

Evaluation of mapping with a tele-operated robot with video feedback.

C. Lundberg, H. I. Christensen

Centre for Autonomous Systems (CAS)

Numerical Analysis and Computer Science, (NADA), KTH

S-100 44 Stockholm, Sweden

+46 73 677 33 95, +46 8 790 6792

carl.lundberg@fhs.mil.se, hic@nada.kth.se

Abstract- This research has examined robot operators' abilities to gain situational awareness while performing tele-operation with video feedback. The research included a user study in which 20 test persons explored and drew a map of a corridor and several rooms, which they had not visited before. Half of the participants did the exploration and mapping using a teleoperated robot (iRobot PackBot) with video feedback but without being able to see or enter the exploration area themselves. The other half fulfilled the task manually by walking through the premises. The two groups were evaluated regarding time consumption and the rendered maps were evaluated concerning error rate and dimensional and logical accuracy. Dimensional accuracy describes the test person's ability to estimate and reproduce dimensions in the map. Logical accuracy handles missed, added, misinterpreted, reversed and inconsistent objects or shapes in the depiction. The evaluation showed that fulfilling the task with the robot on average took 96% longer time and rendered 44% more errors than doing it without the robot. Robot users overestimated dimensions with an average of 16% while non-robot users made an average overestimation of 1%. Further, the robot users had a 69% larger standard deviation in their dimensional estimations and on average made 23% more logical errors during the test.

I. INTRODUCTION

In many of today's robot applications the operator controls or monitors the robot's progress with video feedback from a camera onboard the robot. The performance of these systems largely depends on the operator's ability to achieve Situational Awareness (SA) through the provided interface. According to Endsley [1] SA is defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". The formal definition breaks down into three levels:

Level 1 - Perception of the elements in the environment.

Level 2 - Comprehension of the current situation.

Level 3 - Projection of future status.

The first level includes using senses such as visual, auditory and tactile to gain information about the surrounding environment. On the second level, the disjointed data from level 1 is synthesized, evaluated and prioritized in relation to the present goal to form an understanding of the current state. On the third level, the knowledge from prior levels is used to

predict the oncoming future. A subsequent level depends on the fulfillment of previous. Thus, level 1 needs to be accomplished prior to level 2 and in the same way level 2 has to be reached prior to level 3.

Traditionally research about Situational Awareness has been driven by fields such as power plant control, aircraft, ships, command and control centers for military and large-scale commercial enterprises, intelligence operations, medical systems and information management systems. The SA process occurs, however, not only in the mentioned cases but also is the foundation for decision making in almost every field of endeavor. The elements of importance in gaining SA vary between different fields. But still, the general process of receiving information from the surroundings and filtering, connecting, prioritizing and extrapolating it to predict the future remains the same.

In the case of robot control a fundamental part of the SA will consist of gaining knowledge of the surrounding environment's spatial layout. In this process the operator depends highly on the user-interface while gaining Level 1 and Level 2 SA. It is obvious that it will be harder to gain SA through a robot system compared to being on the spot in person. But how big is the difference? What is the time difference between the two ways and what is the difference in SA accuracy? Questions like these are coming into focus as robot systems that can be put into live missions evolve. Military, police, fire brigade and rescue services are now being forced to evaluate robots as a reconnaissance tool. While doing so they need performance characteristics for comparison of their traditional methods and the robot tools. Performance has to be weighted against costs for acquisition, integration, training and maintenance.

The objective of this research has been to investigate and measure how well an operator gains spatial SA while using a video-feedback robot compared to being there in person. This was done by having two groups of 10 test persons perform a mapping task, one group with and the other without the help of a robot. The two groups were compared regarding time consumption, error rate and accuracy.

II. RELATED WORK

As mentioned the subject of SA has been investigated across wide ranging fields [1]. Also the robotics area has been researched on a variety of subjects. Thus, SA research has



Figure 1. Front view of the PackBot.

been performed on UAVs [10], polymorphic robots [11] and humans in cooperation with autonomous robots [12]. One of the most penetrated area regarding SA in robotics regards scout-robot operation in search and rescue and military settings [2, 3, 4, 5, 6, 7, 9]. This research shows that robot operators have to spend significant time and effort to gain SA in the addressed environments. Information gathering is commonly done by established research methods like interviews, questionnaires, subjective workload measures, observations, communication analysis and ethnographic approaches.

The SA issues are closely coupled to interface design and in many cases the user interface design becomes a driver for evaluation of an operators SA [8, 13, 14].

III. EXPERIMENT DESIGN

A. The robot and the user interface

The iRobot PackBot (Fig. 1) is a portable robot for field use (70*50*20 cm, 18 kg). The battery-powered robot runs on tracks and has a top speed of 3.7 m/s. In standard configuration, the PackBot has a fixed forward facing wide-angle daylight video camera, fixed forward facing IR-camera (Infra Red), IR-illuminator, GPS receiver, electronic compass, and absolute orientation sensors (measuring roll and pitch).

The robot user-interface used during the experiment was based on a rugged laptop, with an external joystick (Fig. 2). During the experiment the user interface was set to only display video from the wide-angle camera with a resolution of 240*320 pixels at a frame rate of 15 frames per second. The wide-angle camera was chosen since its wide angle simplifies manoeuvring through narrow passages, which is especially advantageous for novice operators. In order to simplify the human robot interaction none of the sensor data was displayed and the robot's top speed was limited to 0,7 m/s.

B. Test Persons

The test persons consisted of 12 men and 8 women, who were evenly divided between the robot using group and the non-robot using group. The test persons ranged in age from 24 to 50 years. They all had a college or university education but in varying subjects. The same held true for profession. They



Figure 2. The Laptop user-interface in the experimental setup.

were all frequent computer users but not experienced robot operators or RC-pilots. The test persons had not visited the explored area before.

C. The Explored Region

The test was carried out in a 36 meter long corridor with 15 closed doors and two open doors leading into two shower rooms. The corridor turned 90° at its end, where two temporary walls had been mounted to make the setting more complex (Fig. 3). The explored premises were well lit with fluorescent strip lighting, had no windows and were nearly free from obstacles such as furniture.

The robot-operating group performed their task from a room nearby the explored region. The room was out of sight but within reach of the robot's radio-signal. The non-robot operating group performed the mapping while walking around within the area.

D. Test outline/Course of events:

The test was carried out according to the following steps:

1. Briefing: Informing the participants about the experiment.
2. Prequestionnaire: Handling personal data, experience of robots, tele-operation, joystick control, maps and drawings.
3. Robot training (only for robot operators): In order to ascertain a minimum level of driving competence the robot operators were given a short driving training and had to pass a driving test. The training involved two minutes of driving practice and was followed by a short test along a course similar to the experiment area. The test was repeated until it could be carried through without collisions. All the robot operators passed the test in their first or second try.
4. Map-drawing instructions: In order to simplify evaluation the test subjects were only allowed to draw on the lines on a cross-ruled paper. The scale was set to three checks to a meter and the test persons were instructed to round objects to fit the closest line. Only walls and doors were to be depicted using given symbols (Fig. 4). The test persons were instructed to do the mapping from one end of the area to the other. They were instructed not to return to previously mapped areas unless

those areas led to unexplored regions. The task was to be fulfilled as fast and as accurately as possible. The non-robot operators were instructed to move at normal walking pace. Both groups were instructed to return (themselves or the robot) to the starting point after having covered the whole area.

5. Map-drawing training: The map-drawing training was carried out in order to ascertain that the mapping-instructions were understood and to give the test subjects a chance to practice and ask questions before the start of the real test. The training included exploring and depicting two rooms in the same way as during the experiment. Thereafter, the map was evaluated together with the test leader.

6. Experiment: The mapping task started at the lower end of the corridor (Fig. 3). The maximum time for the mapping was set to one hour, this was, however, not told to the test persons in advance.

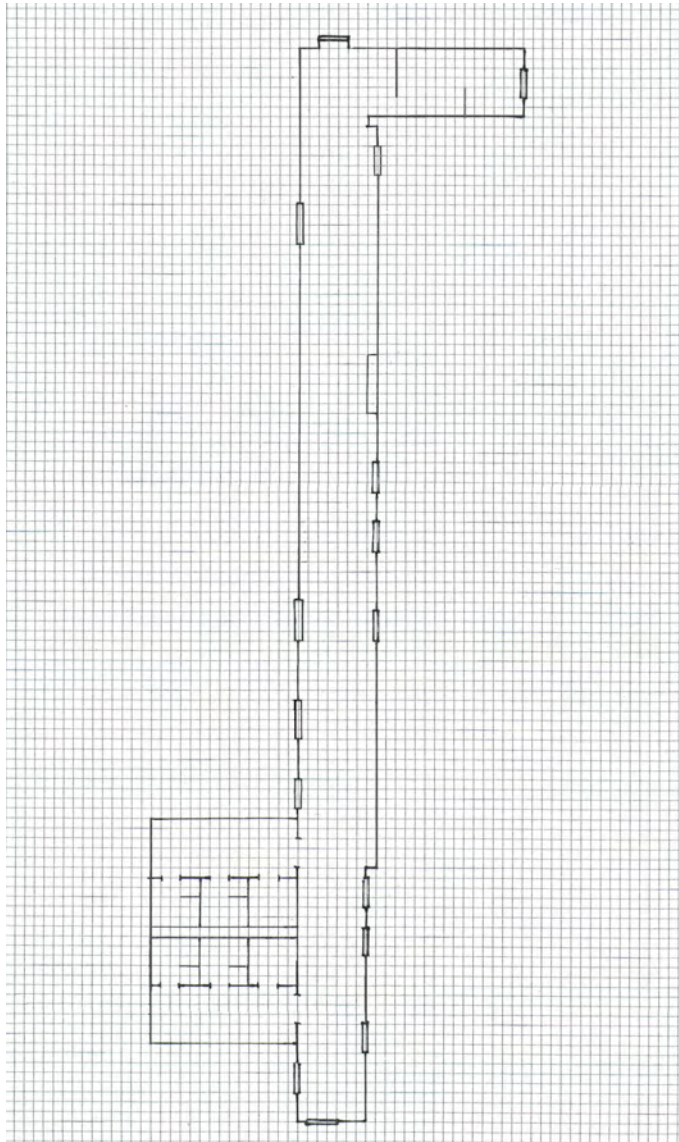


Figure 3. Discretized map of the explored region. The starting point was at the lower end.

7. Postquestionnaire: Handling the test person's experience of the test.

IV. ANALYSIS AND RESULTS

The rendered maps were divided into sub-elements in order to facilitate analysis. Each element had a specified length and was either a wall or a door. The elements start and end where there is a change in element type, in corners or at wall-endings (Fig. 4). The correct version of the discretized map held 112 elements (Fig. 3).

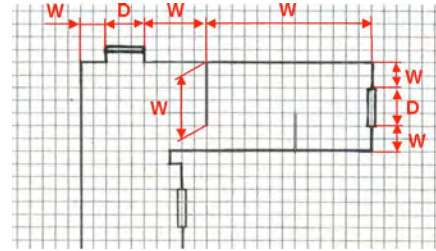


Figure 4. The principle of the dividing the maps into sub-elements. W - wall elements, D - door elements.

The tests were evaluated regarding time consumption, error rate and accuracy. All evaluation criterias were viewed as an average per element. The absolute measures, such as for example total time used, were not applicable for comparison since the test persons did not draw as many elements in their maps. Hence, time consumption was viewed at as the average time in seconds per depicted element. The error rate was regarded as the percentage of erroneous drawn elements.

Errors were divided in two main types, dimensional and logical. Dimensional error was defined as the difference between estimated and true element length expressed as a percentage. The average dimensional errors and the time consumption for the members in the two test groups are displayed in Fig. 5 and Fig. 6.

A test person's dimensional error can be analyzed in two aspects – mean error and standard deviation. The mean error is the average difference between estimated and true element length expressed in percentage. Thus, a constant over or underestimation of element length will render high mean error values. Making the same amount of over as under estimations will, on the other hand, render low mean errors. The standard deviation expresses the consistency of the mean error. A low standard deviation together with a large mean error indicates that the test person made a consistent scaling error such as participant 6 compared to participant 7 in Fig. 6.

The logical errors were grouped into five sub-types:

1. Missing : elements missing, for example a missing wall.
2. Added: elements drawn but not existing in reality.
3. Unexplored: elements not explored due to misinterpretation of the spatial layout. Only one logical error was given for each neglected area although it may have caused several more elements to be missing. This was done since the error was based on one mistake.

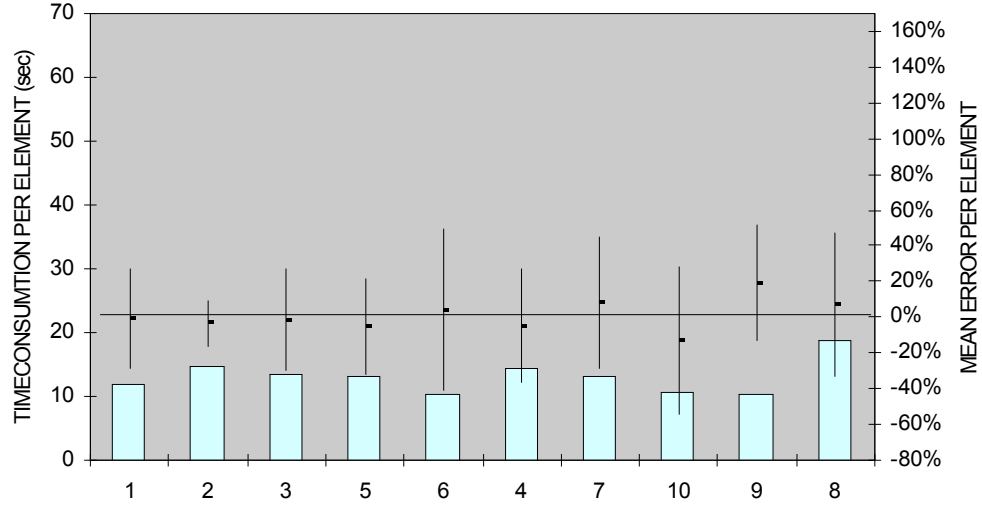


Figure 5. Mean time consumption (bars) and error per element (points with std.dev.) for the 10 non-robot users.

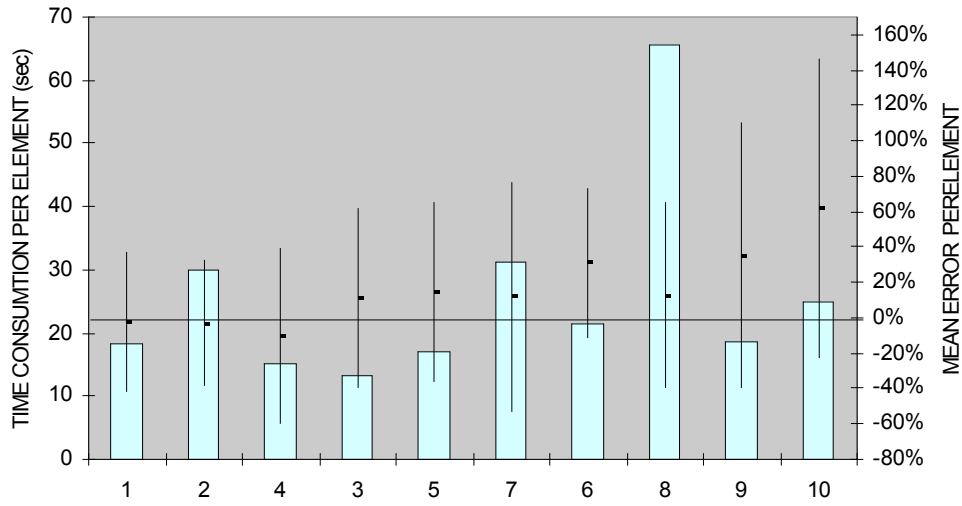


Figure 6. Mean time consumption (bars) and error (points with std.dev.) per element for the 10 robot users.

4. Misshaped: elements with wrong shape for example an element indicating the corridor narrowing instead of widening.

5. Inconsistent: elements whose depictions do not prove consistent from one view to another. For example, a door existing only from one side of a wall.

All logical errors were given a value of one and they were considered compatible enough to be added together in a sum for analysis.

The mentioned performance measures (time consumption, error rate, dimensional mean error, dimensional standard deviation and logical error) were also compiled into an overall performance ranking. This was done by ranking all

the participants compared to each other for the five performance measures and then adding up the individual ranks to an over all rank.

A. Time Consumption

The average time consumption for the robot users and the non-robot users is displayed in Fig. 7 (Fig. 7 displays the average of the time values displayed as bars in Fig. 5 and Fig. 6). On average the non-robot users spent 13 seconds per element while the robot users spent 26. Although the robot-using group took twice as long, it is shown by the high standard deviation and by the individual data in Fig. 6 that some of the robot operators performed as well as some of the non-robot users. This indicates a potential for

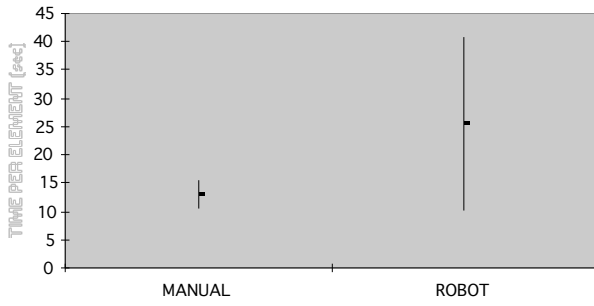


Figure 7. Average time consumption and standard deviation per depicted element for the non-robot users (MANUAL) and the robot users (ROBOT).

improvement depending on factors such as training, talent, motivation, fatigue and experience from fields containing similar mental processing.

B. Error rate

The mean error rate for the two groups, calculated as the percent of elements with a dimensional or logical error, is displayed in Fig. 8. The experiment showed that the non-robot users had an error percentage of 45 % while the robot users had error percentage of 65 %. The two groups had approximately the same standard deviation, 15 for the non-robot users and 12 for the robot users, which indicates a consistent difference in error rate between the two groups.

C. Dimensional error

As displayed in Fig. 9 the non-robot users on average had a mean error of 1 % while the robot users on average had a mean error of 16 % (Fig. 9 displays the average of the mean error values displayed in Fig. 5 and Fig. 6). Hence, the robot operators tended to overestimate dimensions while the non-robot users made approximately as many over as under estimations. Again, the standard deviation implies a larger variation within the robot-using group, which implies potential for improvements such as suggested in the paragraph A. Time Consumption.

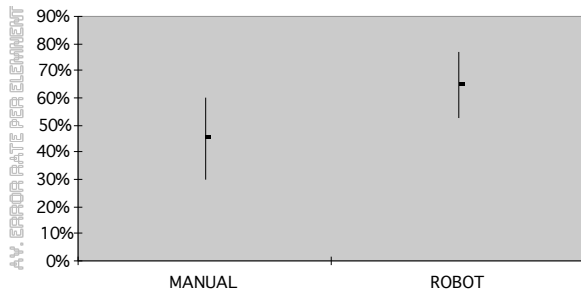


Figure 8. Average error rate in percentages and standard deviation the non-robot users (MANUAL) and the robot users (ROBOT).

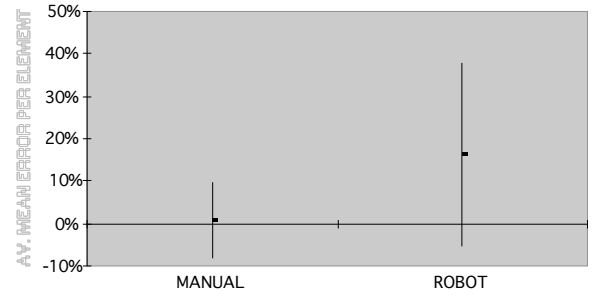


Figure 9. Average mean error in percentage per element for the non-robot users (MANUAL) and the robot users (ROBOT).

The average standard deviation for the two groups, expressing the consistency of the mean error, is displayed in Fig. 10 (Fig. 10 displays the average of the standard deviations displayed in Fig. 5 and Fig. 6). The non-robot group had a mean standard deviation of 32 % while the robot group's corresponding value was 54 %. In this case the standard deviation values (Fig. 10) do not differ significantly, 10 for the non-robot users and 16 for the robot users. This indicates a consistent difference between the two groups, the robot users seem prone to have a greater variation in their dimensional estimations.

D. Logical error

During the test the robot-using group on average made logical errors 4 % of the time depicting an element, Fig. 11. The non-robot users made errors of logic 1.8 % of the time. The standard deviations are alike for the two groups, 1.9 for the non-robot group and 2.2 for the robot group. This again implies that there is a consistent difference between the two groups regarding logical errors.

E. Over all ranking

The rankings for the five performance criteria displayed in Table I, show that the robot users are generally overrepresented in the lower end compared to the non-robot users. The robot operators occasionally manage to compete

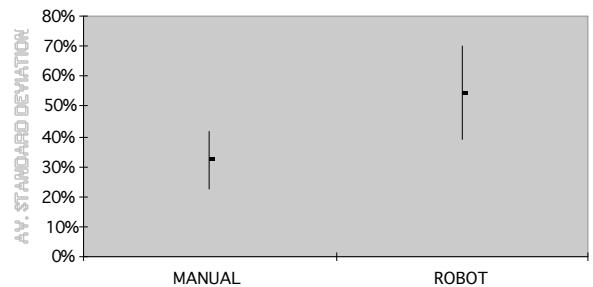


Figure 10. Average standard deviation of dimensional mean error for the non-robot users (MANUAL) and the robot users (ROBOT).

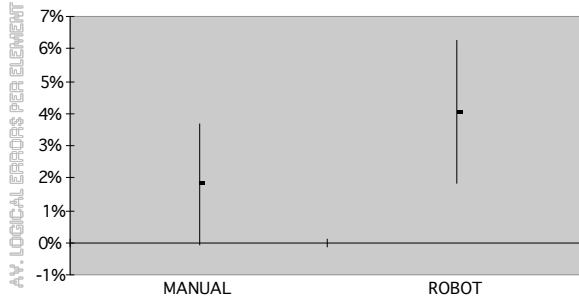


Figure 11. Average logical errors and standard deviation per element for the non-robot users (MANUAL) and the robot users (ROBOT).

with the non-robot users like in the case of dimensional mean error. However, in the total rank the non-robot operators predominate the higher ranks.

According to the questionnaires the two best performing robot operators were both highly experienced in the interpretation of 3-dimensional computer representations.

ROBOT 1: Male, 37 years, Industrial designer and product developer, professional 3D-CAD-user, daylily computer user with medium skill, seldom or never plays computer games, has tried to operate RC-crafts a few times, inexperienced to joystick control, inexperienced to robot operation.

ROBOT 2: Female, 29 years, MSc. of Ergonomic design and Production, professional 3D-CAD-user, daylily computer user with medium skill, seldom or never plays computer games, has tried to operate RC-crafts a few times, inexperienced to joystick control, inexperienced to robot operation.

V. DISCUSSION

There are a number of factors to consider when analyzing a persons robot aided exploration of a building. Which factors are most interesting will vary with the purpose of the mission. In some cases it might be most important to search for certain objects. In other cases it might be of greatest interest to find a passage through a building and in yet another, the purpose might be to cover most possible area. Similarly, the impact from different types of errors might vary between missions. For example, in some cases it might not matter if the dimensions are accurate as long as the logics are correct. No matter the purpose of the mission it will be of interest to do some sort of navigation which includes creating a mental model of the spatial layout. In this experiment the operators were forced to draw a map during the exploration that in it self is a violation of the spontaneous way the operator might have approached the task in a real case. It is reasonable to believe that the drawing process made the exploration more time consuming. Further, the demand to draw a map probably improved the accuracy of the spatial mental model since it forced the operator to mentally process

TABLE I
ALL PARTICIPANTS ORDERED ACCORDING TO RANK IN THE DIFFERENT PERFORMANCE CRITERIAS AND FOR THE TOTAL RANK. WITH THE BEST PERFORMERS LISTED AT THE TOP. THE NON-ROBOT USERS ARE NAMED MANUAL 1-10. THE ROBOT USERS ARE NAMED ROBOT 1-10 AND SHADED.

TIME CONSUMPTION	ERROR RATE	DIMENSIONAL DIFFERENCE	DIMENSIONAL STANDARD DEVIATION	LOGICAL ERROR	TOTAL RANK
MANUAL 6	MANUAL 2	MANUAL 1	MANUAL 2	MANUAL 1	MANUAL 1
MANUAL 9	MANUAL 1	MANUAL 3	MANUAL 5	MANUAL 4	MANUAL 2
MANUAL 10	MANUAL 5	ROBOT 1	MANUAL 1	MANUAL 6	MANUAL 3
MANUAL 1	MANUAL 3	ROBOT 2	MANUAL 3	MANUAL 6	MANUAL 4
MANUAL 7	ROBOT 2	MANUAL 2	MANUAL 4	ROBOT 5	MANUAL 5
MANUAL 5	MANUAL 4	MANUAL 6	MANUAL 9	ROBOT 6	MANUAL 6
ROBOT 3	MANUAL 7	MANUAL 4	ROBOT 2	MANUAL 2	MANUAL 7
MANUAL 3	ROBOT 1	MANUAL 5	MANUAL 7	MANUAL 3	ROBOT 1
MANUAL 4	ROBOT 3	MANUAL 8	ROBOT 1	MANUAL 3	MANUAL 8
MANUAL 2	MANUAL 6	MANUAL 7	MANUAL 8	MANUAL 7	ROBOT 2
ROBOT 4	MANUAL 10	ROBOT 4	MANUAL 10	MANUAL 5	MANUAL 9
ROBOT 5	MANUAL 8	ROBOT 3	ROBOT 6	MANUAL 9	MANUAL 10
ROBOT 1	ROBOT 4	ROBOT 7	MANUAL 6	ROBOT 4	ROBOT 3
ROBOT 9	MANUAL 9	ROBOT 8	ROBOT 4	ROBOT 3	ROBOT 4
MANUAL 8	ROBOT 8	MANUAL 10	ROBOT 5	MANUAL 10	ROBOT 5
ROBOT 6	ROBOT 7	ROBOT 5	ROBOT 3	ROBOT 2	ROBOT 6
ROBOT 10	ROBOT 5	MANUAL 9	ROBOT 8	ROBOT 7	ROBOT 7
ROBOT 2	ROBOT 6	ROBOT 6	ROBOT 7	ROBOT 9	ROBOT 8
ROBOT 7	ROBOT 9	ROBOT 9	ROBOT 9	ROBOT 10	ROBOT 9
ROBOT 8	ROBOT 10	ROBOT 10	ROBOT 10	ROBOT 8	ROBOT 10

the acquired information. The map was probably also a significant memory-support for the test-persons during the exploration. The restriction to the cross-ruled paper might have influenced the test-persons to be more structured in their map drawing. The cross-rules also prevented depiction of any curved shapes (there were no curved walls in the explored region). Earlier test indicate that robot operators have more trouble with curved than straight walls.

The robot imposes a number of perceptual disadvantages to the operator. The wide-angle lens makes driving easier but it also distorts the perspective, which makes recognizing objects and judging dimensions harder. The resolution of 240*320 pixels is significantly lower than the resolution of the human eye. The floor level placement of the camera gives an unusual perspective and obstacles also easily block it. In addition to drawbacks in visual feedback the robot also lacks inertial and motory information of movements. Despite mentioned disadvantages, the operation of the robot system did not prove too difficult for the test persons. Even with low training level they managed to get around with a reasonable number of collisions (mainly when passing through doorways). The non-robot users did not move at a much higher pace than the robot during the exploration, most of the time was spent viewing and drawing.

Regarding analysis, it is not obvious in what way the different performance criteria should be valued towards each other. As mentioned it will largely depend on the purpose of the mission. The chosen evaluation strategy emphasizes dimensional errors done on small elements (gives a higher error percentage). Further the logical errors were all valued the same although the consequence of the may vary depending on mission type.

The general validity of gained results is influenced by a number of factors. The robot operators had a minimum of robot training. The performances by the better robot operators indicate a potential for general improvement through training. The experiment was done in a fairly simple and uncluttered environment of a type familiar to the test persons with good light conditions. Earlier studies

indicate that environments with less familiar objects or very cluttered environments complicate robot-aided exploration.

The results are likely to be valid for systems with similar fields of view, resolution and camera placement. New technology like omni directional cameras, stereovision or support from other sensor types will probably have major performance impact.

VI. CONCLUSIONS

The result of this research describe how well an operator can gain spatial knowledge while performing exploration with a video-feedback robot compared to being in the place in person. The robot-operating group needed 69 % more time and were wrong 44 % more often than the non robot users during the exploration. Robot users overestimated dimensions with an average of 16% while non-robot users only made an average overestimation of 1 %. Further, the robot users on average had a 69 % larger standard deviation in their dimensional estimations and on average made 23 % more logical errors during the test. However, it is indicated by high performing test participants that it might be possible to decrease time consumption and mean error.

The results can be used as a guideline for the performance of other systems. The validity is influenced by factors of three kinds:

1. Operator – The test participants were arbitrary chosen novices. Training, talent, motivation, fatigue and skills in fields containing similar mental processing are probable to have influence on the operator's performance.

2. Environment - While these tests were carried out in an environment well suited for robot exploration it is known that there is a strong relation between environmental complexity and the prospect of gaining SA through robot exploration.

3. Robot – The robot used during the experiment had a minimum of features. Interface design, camera performance and placement, maneuverability, user interface design as well as integration of other sensors influences the system efficiency.

REFERENCES

- [1] M.R. Endsly, B Bolté, D.G. Jones, “*Designing for Situation Awareness – An Approach to User-Centered Design*”, Taylor & Francis 2003, ISBN 0-748-40966-1, pp13-30.
- [2] Yanco, H. A. and Drury, J. L. (2004). “‘Where am I?’ Acquiring situation awareness using a remote robot platform.” Proceedings of the IEEE Conference on Systems, Man and Cybernetics, The Hague, Netherlands, October.
- [3] Burke, J.L., Murphy, R.R., Coover, M. D., and Riddle, D.L.(2004). “*Moonlight in Miami: A Field Study of Human-Robot Interaction in the Context of an Urban Search and Rescue Disaster Training Exercise*.” Human-Computer Interaction. 19(1-2), pp. 85 – 116.
- [4] Casper, J. (2002). “*Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center*.” MS Thesis, University of South Florida Department of Computer Science and Engineering.
- [5] Casper, J.L., Micire, M.J. and Li Gang, R. (2004). “*Inuklut Services Ltd. – Search and Rescue Robotics*.” Proc. 3rd International Conference on Continental Earthquakes (ICCE), Beijing, China.
- [6] Casper, J., and Murphy, R.R. (2003). “*Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center*.” IEEE Transactions on Systems, Man, and Cybernetics, Part B, June, vol. 33, pp. 367-385.
- [7] Drury, J. L., Scholtz, J., and Yanco, H. A. (2003). “*Awareness in human-robot interactions*.” Proceedings of the IEEE Conference on Systems, Man and Cybernetics, Washington, DC, October.
- [8] Keskinpala H., and Adams J., “*Objective Data Analysis for a PDABased Human-Robotic Interface*”, 0-7803-8566-7/04, 2004 IEEE.
- [9] C. Lundberg, H.I. Christensen, A. Hedsrom, “*The Use of Robots in Harsh and Unstructured Field Applications*” Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication, 2005.
- [10] J.L. Dury, L Riek, N. Rackliffe “*A Decomposition of UAV-Related Situational Awareness*”, Proceedings of the ACM Conference on Human-Robot Interaction, 2006.
- [11] J. L. Drury, H. A. Yanco, W. Howell, B. Minten, J. Casper “*Changing Shape: Improving Situation Awareness for a Polymorphic Robot*”, Proceedings of the ACM Conference on Human-Robot Interaction, 2006.
- [12] B. P. Sellner, L. M. Hiatt, R. Simmons, S. Singh, “*Attaining Situational Awareness for Sliding Autonomy*”, Proceedings of the ACM Conference on Human-Robot Interaction, 2006.
- [13] C. W. Nielsen, M. A. Goodrich, “*Comparing the Usefulness of Video and Map Information in Navigation Tasks*”, Proceedings of the ACM Conference on Human-Robot Interaction, 2006.
- [14] T. W. Fong, C. Thorpe, and C. Baur. “*Advanced interfaces for vehicle teleoperation: Collaborative control, sensor fusion displays, and remote driving tools*.” Autonomous Robots, 11(1):77–85, July 2001.