Human Based Benchmark for Robot Navigation Assessment

Eduardo O. Freire #1, Elyson A. N. Carvalho #2, Marcos V. S. Alves #3, Jugurta Montalvão #4, Lucas Molina #5

Electrical Engineering Department of Federal University of Sergipe – UFS

Av. Marechal Rondon, S/N, São Cristóvão-SE, CEP 49100-000, Brazil

1 efreire@ufs.br

2 ecarvalho@ufs.br

3 marcoxvini88@yahoo.com.br

4 jmontalvao@ufs.br

5 lmolina@ufs.br

***Abstract*—This paper proposes the use of a human based benchmark for mobile robot navigation architectures. In order to do so, a simulator was developed to allow the realization of experiments in which both humans and automatic controllers should perform the very same task using the same mobile platform (with the same holonomic restrictions) and using the same sensorial information, considering that there is no previous knowledge about the environment. Four hypotheses were proposed and some experiments were performed to evaluate them. The obtained results are presented and discussed.**

I. INTRODUCTION

With the increasing development of mobile robotics and increasing number of works in which new navigation architectures are proposed, the need for standardization of methods for assessing and comparing the performance of these architectures is a reality [3], [4], [5]. Therefore, several methodologies had been proposed as possible standards or benchmarks [4], [10]. Such methodologies attempt to standardize and describe in details aspects such as the assessment procedure, the specifications of the mobile platform used to perform the experiments, and the parameters or metrics used.

In most of the cases, assessment is made based on metrics such as elapsed time, traveled distance to reach the destination point, mean linear velocity, trajectory smoothness, etc. The optimal value for each metric is obtained from what should be an "optimal trajectory". In fact, such "optimal trajectory" is just an abstraction of the authors. Besides, considering the robot characteristics and the available information, such trajectory may be impossible to be attained by the robot.

There is a very hard question to answer: the trajectory is, or must be optimal in which sense? Which feature should be prioritized? Velocity? Security? Elapsed time? Smoothness? Energy consumption? A possible answer is that it depends on the application, and it is true. Even though, this is a controversial and open question.

On the other hand, some examples can be found in literature about the use of human performance or standards to assess the performance of different kinds of algorithms. The mp3 audio compression standard is a remarkable example of this [9], [12].

Previous codification methods were based on models of the signal source. Instead of this, the codification method in which the mp3 standard is based on was developed considering the model of the human auditory system [9]. The high performance of this compression algorithm is possible because it takes advantage of auditory masking effect. This masking is a perceptual failure of the human ear that occurs when the presence of a strong audio signal masks a spectral neighborhood of weaker audio signals [12].

This phenomenon was observed and confirmed through a variety of psychoacoustic experiments [14]. The developed model became known as the psychoacoustic model [12], and the resultant audio compression standard is one of the most accepted and widespread in the world and one of the reasons that enabled the success of this method was the use of human hearing ability as a quality standard.

Similarly, humans have a native ability to locate and move around in various types of environment. Thus, this work proposes the use of a human based benchmark for mobile robot navigation architectures. Using a simulator, especially designed for this purpose, the performance of already tested and published navigation architecture, the *Tangential Escape* [6], was compared with the performance of a group of persons. To allow a fair comparison, both, humans and the navigation architecture, were submitted to the same holonomic restrictions and using the same sensorial information.

The *Tangential Escape* Architecture will not be explained in this paper. Its main idea consists in temporarily changing the position of the destination point when an obstacle is detected so the robot follows a tangential trajectory with respect to the detected obstacle, and returning to the real destination point when the obstacle is avoided. This architecture was chosen because even though its simplicity, it is capable of attain a good performance in complex environments. A detailed description of the *Tangential Escape* architecture can be found in [6].

The paper is organized as follows. In Section II, the simulator is described. The proposed methodology is presented in Section III. Section IV is dedicated to the obtained results and the discussion about them. Finally, conclusions and future work are presented in Section V.

II. THE SIMULATOR

In order to allow a fair comparison between the performance of the navigation architecture and the human performance, it would be necessary to find a way to force the human to guide the robot under the same holonomic restrictions and based on the same limited sensorial information as the navigation architecture.

To solve this problem, a simulator was especially designed and implemented. The simulated robot may be manually controlled by a human or automatically controlled by the implemented navigation architecture. In both cases, the holonomic restrictions and the available sensory information are the same.

The simulated robot uses differential traction (most widely used). It can be manually controlled using the keyboard arrows. While the up arrow is pressed, the robot moves with a 300 mm/s linear velocity. The left and right arrows are used to control the angular velocity. The left arrow commands a 0.5 rad/s anti-clockwise angular velocity, while the right arrow commands a 0.5 rad/s clockwise angular velocity. Such velocities, besides been compatible with real applications [6],[7], were preferred by the vast majority of volunteers in the preliminary stages of simulator test and therefore were chosen. The angular and linear velocities were kept constant because the volunteers felt that, in this way, the robot guiding task was easier. In future versions of the simulator, the possibility of changing linear and angular velocities will be re-introduced, however, using new control methods, like a joystick or a touch pad. Moreover, the down arrow, which initially was used to command a negative linear velocity, was disabled, since the impossibility to move backwards is a common assumption in the control of mobile robots. The *Tangential Escape* architecture was also implemented with this same restriction.

The simulated sensory system is similar to a range sensor capable of measure distances around the robot (360 degrees) with a resolution of 1 degree, like a laser range sensor.

As the simulated sensors are similar to range sensors, obviously they cannot provide information about parts of the environment that are behind walls, closed doors or other obstacles. All of them were assumed to be opaque.

The range of measurement of the simulated sensory system may be selected by the user. In the simulations performed in this work a short range, about 1.15 meter, was selected. The reason for this is because most of the obstacle avoidance algorithms only actuate when obstacles are close to the robot, about 1 meter, for example. In the case of the *Tangential Escape* architecture here implemented, the distance considered for obstacle avoidance action is 0.75 meter. By doing so, the idea was again to guarantee that the volunteers should be submitted to the same conditions imposed to the navigation architecture, aiming in to allow a fair comparison. Thus, the volunteers were able to see just the part of the environment covered by the simulated sensory system, and as the robot moves in the environment, new parts of it are revealed, while others disappear, as can be seen in Fig. 1.

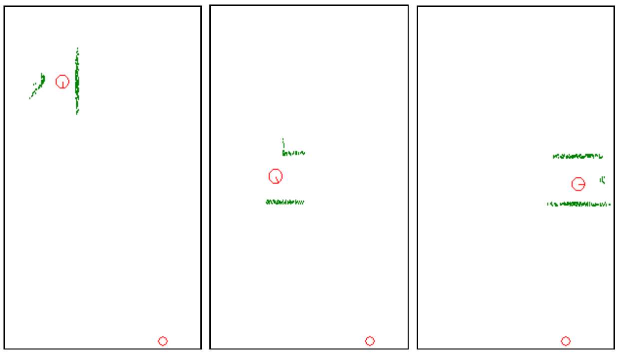

Fig. 1. Image sequence obtained during the realization of an experiment

In a preliminary version of this work [16] was not possible to simulate the measurement noise. With the current version of the simulator it is possible to choose to perform the simulations with or without noise (Gaussian). The noise level may also be selected.

In Fig. 1 a sequence of three images obtained during the realization of an experiment is shown. It is possible to see the robot, represented by a red circle and a straight line indicating its front side and its orientation. The robot dimensions were selected to correspond to a robot whose base diameter is about 40 cm. Also, the portion of the environment covered by the sensory system is shown.

Another red circle that appears in the bottom part of the image represents the destination point. Notice that it is not possible to know if there is any obstacle between the current robot position and its destination point.

The environment map can be selected by the user. Any bitmap image can be used. Obstacles or environment limits must be represented by black pixels, while free space is represented by white pixels. So, environments can be created or modified using any software capable to work with bitmap image files.

III. METHODOLOGY

The purpose of this work may be stated in the form of the following hypotheses:

A. Human navigation ability may be used as a benchmark for robot navigation architectures;
B. Noise should affect navigation architectures more than humans;
C. Humans are benefited by their memory capability;
D. For unknown environments (no memory effects) navigation architectures are more efficient (considering the proposed metrics) but humans have a greater capacity of generalization.

All tests were performed using the same software and hardware. The software was the simulator described in Section II, and the hardware was a notebook with a standard US International keyboard and with the following configuration: Core 2 Duo 2.5GHz Intel processor, 6 MB cache memory and 4 GB DDR2 667MHz memory.

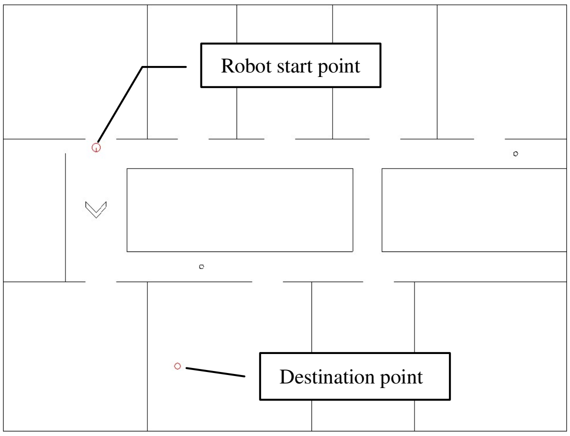

Fig. 2. First task

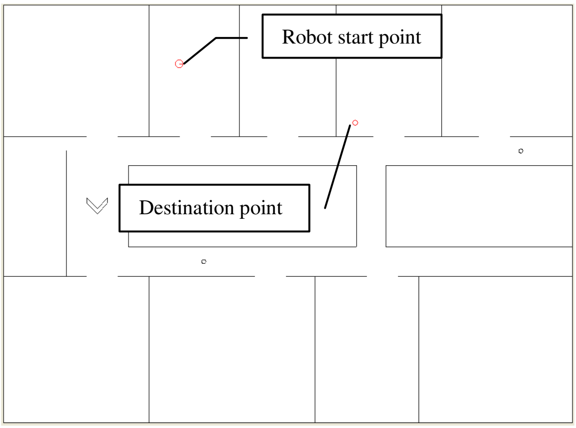

Fig. 3. Second task

Before the execution of the experiments, each volunteer was briefed about the correct use of the simulator and trained in three preliminary tasks, performed in three different training environments. The volunteers were allowed to train until they felt prepared to the real experiments. The training environments were a free space environment, a labyrinth without other obstacles than the walls, and an environment with rooms, walls, obstacles and local minima points. This training stage was of fundamental importance to allow the volunteers to acquire the necessary skills to manually control the robot, without affecting the experiment performance.

Moreover, among the volunteers, one of them was submitted to a much more exhaustive training stage, thus developing excellent skills to manually control the simulated robot. This volunteer will be referred as the "specialist".

Two tasks were considered. The first one consists in the map presented in Fig. 2, and in order to check the proposed hypotheses the following steps were performed:

1. Volunteers executed the task without measurement noise and with no information about the environment.
2. The task was repeated (volunteers already have some information about the environment due to memory), but this time with the introduction of measurement noise.
3. The task is repeated again, with noise, but this time the volunteers can see the entire map of the environment.
4. The task is repeated one more time, with no measurement noise and without the view of the environment map. This step is similar to the first one, but this time the volunteers already saw the entire environment map and should remember it.

The second task was designed to check hypothesis B (noise should affect navigation architectures more than humans). In order to do so, the volunteers could not have any information about the environment structure. So, the task took place in a distinct area of the environment, as shown in Fig. 3. Also, the second task was performed before the volunteers know the environment map (before step #3 of the first task). As previously mentioned this second task was performed without the view of the environment map, and with the presence of measurement noise.

The total number of volunteers was 12. This universe is relatively restricted and do not allow to reach definitive conclusions. However, this stage of the work still has the objective to raise the question of using a human benchmark for mobile robot navigation assessment, through the proposition of four hypotheses and the presentation of several simulated results, which provided some important evidences about the proposed approach. Besides, the work also has the objective to introduce the proposed simulator.

IV. Results and Discussion

The results corresponding to the four steps of the first task and to the second task are presented. In each case a table and a graphic are presented (disposed side by side). The metric values for each volunteer and for the *Tangential Escape* architecture are presented in Tables I to V. Moreover, the collection of trajectories followed by each volunteer and by the *Tangential Escape* architecture is presented in Fig. 4 to 8.

The analysis of Table I and Fig. 4 (corresponding to step #1 of the first task – experiment performed without measurement noise or previous information about the environment) reveals that when humans have no previous information about the environment (no memory) and in the absence of measurement noise, the *Tangential Escape* architecture is the winner considering all proposed metrics.

When considering the presence of measurement noise the performance of the *Tangential Escape* architecture is strongly affected (see Table II and Fig. 5). On the other hand, all human's metrics improved after they acquired some knowledge about the environment, even in presence of noise. It is important to notice that the *Tangential Escape* architecture does not have memory capability.

When the entire environment map is visible for the volunteers, human results presented in Table III are even better than those presented in Table II. The same conclusion can be reached comparing Fig. 6 and Fig. 5 (except for one of the volunteers whose trajectory presented in Fig. 6 is very distinct from the others). These results correspond to step #3 of the first task. In this case, the results of the *Tangential Escape* architecture are the same as in Table II, since this approach is not able to take advantage of the knowledge about the environment map.

TABLE I
METRICS USED TO PERFORMANCE ASSESSMENT: STEP #1 OF THE FIRST TASK

Operator	Traveled Distance (mm)	Time (s)	Mean Linear Velocity (mm/s)	Smoothness (degrees)
Tangential Escape	18347	69.5	264	0.71
Specialist	18894	72.3	261	2.65
Volunteer 1	19119	87.2	219	3.91
Volunteer 2	21544	112.0	192	2.00
Volunteer 3	20022	78.9	254	2.61
Volunteer 4	COLISION			
Volunteer 5	20022	76.1	263	3.68
Volunteer 6	20924	161.3	130	1.88
Volunteer 7	19261	72.0	268	2.74
Volunteer 8	19599	87.1	225	3.92
Volunteer 9	20445	91.1	224	3.08
Volunteer 10	19825	85.4	232	2.65
Volunteer 11	21799	88.3	247	5.70
Mean value (Humans)	**21800**	**98.3**	**253**	**3.23**

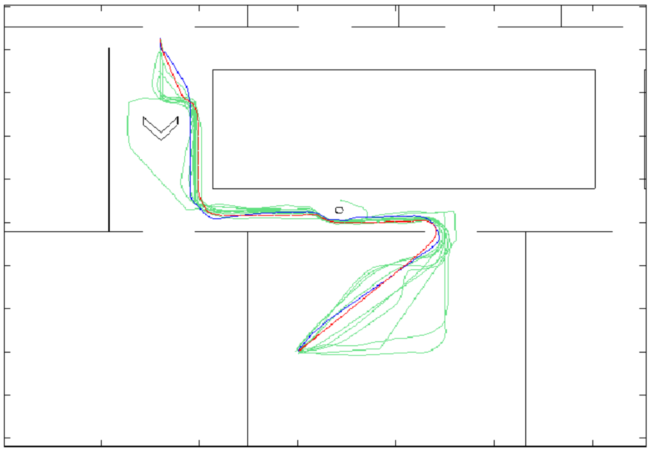

Fig. 4. Step #1 of the first task: trajectory followed by the Tangential Escape architecture (red), trajectory followed by the specialist (blue) and similar trajectories followed by the volunteers (green).

TABLE II
METRICS USED TO PERFORMANCE ASSESSMENT: STEP #2 OF THE FIRST TASK

Operator	Traveled Distance (mm)	Time (s)	Mean Linear Velocity (mm/s)	Smoothness (degrees)
Tangential Escape	18834	107.7	175	0.69
Specialist	19655	76.9	256	2.70
Volunteer 1	19317	71.4	271	1.50
Volunteer 2	20840	82.3	253	2.30
Volunteer 3	18838	67.0	281	2.69
Volunteer 4	COLISION			
Volunteer 5	20699	82.8	250	2.69
Volunteer 6	20473	150.3	136	2.61
Volunteer 7	18922	72.2	262	4.58
Volunteer 8	18922	78.3	242	2.35
Volunteer 9	19345	71.9	269	2.70
Volunteer 10	19684	79.5	248	2.62
Volunteer 11	21799	86.7	251	4.14
Mean value (Humans)	**19863**	**83.6**	**247**	**2.81**

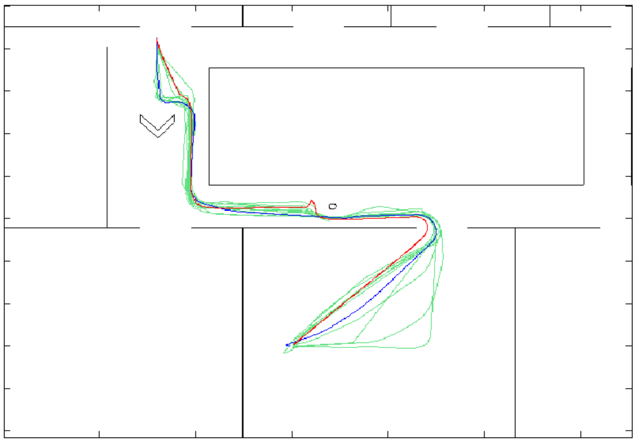

Fig. 5. Step #2 of the first task: trajectory followed by the Tangential Escape architecture (red), trajectory followed by the specialist (blue) and similar trajectories followed by the volunteers (green).

TABLE III
METRICS USED TO PERFORMANCE ASSESSMENT: STEP #3 OF THE FIRST TASK

Operator	Traveled Distance (mm)	Time (s)	Mean Linear Velocity (mm/s)	Smoothness (degrees)
Tangential Escape	18834	107.7	175	0.69
Specialist	18697	66.1	283	1.57
Volunteer 1	18838	67.4	279	2.77
Volunteer 2	19881	72.6	274	1.57
Volunteer 3	18781	66.4	283	2.68
Volunteer 4	18274	64.6	283	1.75
Volunteer 5	18640	65.9	283	2.72
Volunteer 6	20699	121.8	170	2.77
Volunteer 7	18866	66.7	283	2.78
Volunteer 8	18640	65.9	283	1.63
Volunteer 9	18386	65.2	282	1.59
Volunteer 10	18415	71.9	256	1.72
Volunteer 11	18866	66.9	282	4.04
Mean value (Humans)	**18915**	**71.8**	**270**	**2.30**

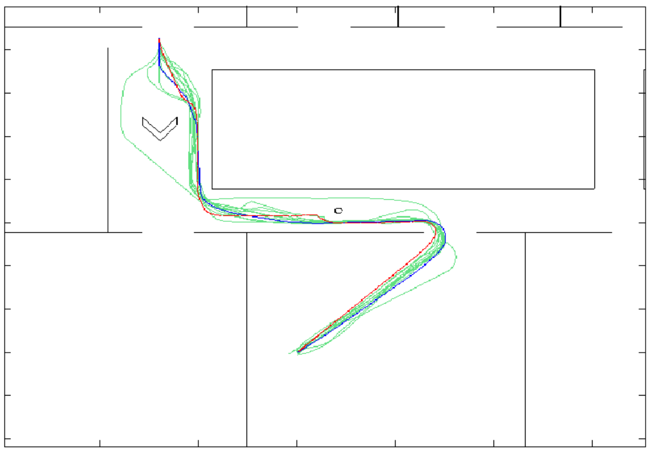

Fig. 6. Step #3 of first task: trajectory followed by the Tangential Escape architecture (red), trajectory followed by the specialist (blue) and similar trajectories followed by the volunteers (green).

TABLE IV
METRICS USED TO PERFORMANCE ASSESSMENT: STEP #4 OF THE FIRST TASK

Operator	Traveled Distance (mm)	Time (s)	Mean Linear Velocity (mm/s)	Smoothness (degrees)
Tangential Escape	18347	69.5	264	0.71
Specialist	19120	68.1	281	2.74
Volunteer 1	19345	72.1	268	2.69
Volunteer 2	19796	75.5	262	1.59
Volunteer 3	18866	67.0	282	3.92
Volunteer 4	19430	72.1	269	3.71
Volunteer 5	19091	67.5	290	1.68
Volunteer 6	19796	109.9	180	2.02
Volunteer 7	18640	66.6	280	3.93
Volunteer 8	19091	68.4	279	2.74
Volunteer 9	19599	77.4	253	2.58
Volunteer 10	18838	66.9	282	4.90
Volunteer 11	19430	70.1	277	3.99
Mean value (Humans)	**19253**	**73.5**	**267**	**3.04**

TABLE V
METRICS USED TO PERFORMANCE ASSESSMENT: SECOND TASK

Operator	Traveled Distance (mm)	Time (s)	Mean Linear Velocity (mm/s)	Smoothness (degrees)
Tangential Escape	17348	102.4	169	0.62
Specialist	18499	66.6	278	1.39
Volunteer 1	31781	129.9	245	2.16
Volunteer 2	19007	77.4	246	2.48
Volunteer 3	39818	154.3	258	1.91
Volunteer 4	29328	119.7	245	3.62
Volunteer 5	33276	124.2	268	2.69
Volunteer 6	39283	262.5	150	2.24
Volunteer 7	17710	64.1	276	2.66
Volunteer 8	41454	166.5	249	3.78
Volunteer 9	17907	68.0	263	1.33
Volunteer 10	17879	67.6	264	3.74
Volunteer 11	32825	123.8	265	5.18
Mean value (Human)	**28231**	**118.7**	**251**	**2.76**

Finally, the last step of the first task (see Table IV and Fig. 7) is the repetition of the first step, what means that there is no measurement noise and the environment map is hidden. So, the results of the *Tangential Escape* architecture are the same as in Table I, since this approach does not have memory capability. On the other hand, humans are benefited by the previous knowledge acquired in previous simulations. This conclusion can be reached by comparing data from Table IV and Table I. When compared with data from Table III the results shown in Table IV are worst. When the entire map is visible, humans perform better than using only the available sensory data and their memory capability.

The data presented in Table V should be analysed comparing the results obtained by humans against those obtained by the *Tangential Escape* architecture, both considering the presence of measurement noise and no previous knowledge about the environment (also see Fig. 8).

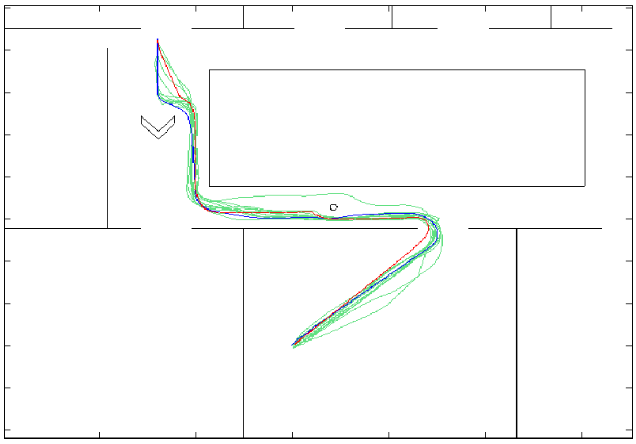

Fig. 7. Step #4 of the first task: trajectory followed by the Tangential Escape architecture (red), trajectory followed by the specialist (blue) and similar trajectories followed by the volunteers (green).

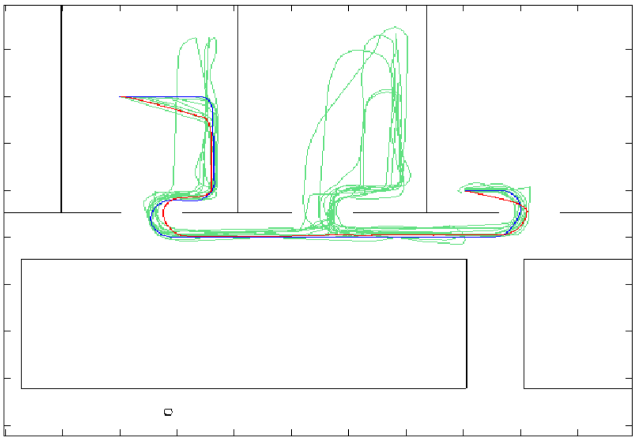

Fig. 8. Second task: trajectory followed by the Tangential Escape architecture (red), trajectory followed by the specialist (blue) and similar trajectories followed by the volunteers (green).

As can be noticed analysing Table V, when noise was not considered, the *Tangential Escape* architecture performance do not always surpasses humans. Considering smoothness and travelled distance the *Tangential Escape* architecture is still competitive, but when considering elapsed time and mean linear velocity (which are dependent characteristics), the *Tangential Escape* architecture is surpassed by five volunteers.

V. CONCLUSIONS

The obtained results suggest that human navigation ability in fact can be used as a benchmark for robot navigation architectures –hypothesis A.

The trajectories followed by the majority of the volunteers are really similar to the trajectory followed by the *Tangential Escape* architecture. Also, in the absence of noise and previous knowledge about the environment, the *Tangential Escape* architecture was unbeatable.

Hypothesis B is corroborated by the results of the second task (Table V and Fig. 8), in which no previous knowledge about the environment was available, but measurement data was contaminated by noise. The *Tangential Escape* architecture performance degradation was much more significant then when considering human's results. This hypothesis is also corroborated by data obtained in step #2 of the first task, when compared with data obtained in step #1 of the same task. Considering only the data related with the *Tangential Escape* architecture, it is obvious that the performance of the *Tangential Escape* architecture in step #1 was much better than the performance attained in step #2, due to the presence of noise. So, that is strong evidence that navigation architectures are more severely affected by noise than humans.

Hypothesis C is corroborated by the comparison of the results obtained by humans in steps 1 and 4 of the first task. The results presented in Table IV and Fig. 7, obtained by the volunteers after performing the same task some times and after know the entire environment map, were much better than those presented in Table I and Fig. 4, obtained when the volunteers had no previous knowledge about the environment.

Finally, despite of some important evidences, the available data is not enough to definitively prove the proposed hypotheses. This conclusion is especially true when considering hypothesis D. In this case it is even unclear if obtained results corroborate this hypothesis. Further studies will be held to attempt to prove the proposed hypotheses.

ACKNOWLEDGMENT

The authors wish to thanks CENPES/PETROBRAS and CNPq for financial support and the award of fellowships during investigation period, and the volunteers that kindly performed the proposed experiments.

REFERENCES

[1] J. S. Albus, "Metrics and Performance Measures for Intelligent Unmanned Ground Vehicles," in *Proc. of the 2002 Performance Metrics for Intelligent Systems Workshop*, 2002.

[2] F. Amigoni, "Experimental evaluation of some exploration strategies for mobile robots," in *Proc. ICRA'08*, 2008, p. 2818–2823.

[3] F.P. Bonsignorio, J. Hallam and A. P. del Pobil, "Good Experimental Methodologies in Robotics: State of the Art and Perspectives," in *Proc. of the Workshop on Performance Evaluation and Benchmarking for Intelligent Robots and Systems, IEEE/RSJ International Conference on Intelligent Robots and Systems,* 2007.

[4] D. Calisi, L. Iocchi, and D. Nardi, "A unified benchmark framework for autonomous Mobile robots and Vehicles Motion Algorithms (MoVeMA benchmarks)," in *Proc.* of the *RSS Workshop on Experimental Methodology and Benchmarking*, 2008.

[5] N. D. M. Ceballos, N. J. L. Ospina and J. A. Valencia, "Performance Metrics for Robot navigation," in *Proc. of the Electronics, Robotics, And Automotive Mechanics Conference,* 2007, p. 749.

[6] A. Ferreira, F. G. Pereira, R. F. Vassallo, T. F. Bastos-Filho; M. Sarcinelli-Filho, "An approach to avoid obstacles in mobile robot navigation: the *Tangential Escape*," SBA. Sociedade Brasileira de Automática, v. 19, pp. 395-405, 2008.

[7] E. O. Freire, T. F. Bastos-Filho, M. Sarcinelli-Filho and R. Carelli, "A New Mobile Robot Control Approach via Fusion of Control Signals". IEEE Transactions on Systems, Man and Cybernetics. Part B, Cybernetics, v. 34, pp. 419-429, 2004.

[8] M. Irie, K. Nagatani and A. Gofuku, "Path evaluation for a mobile robot based on a risk of colision," in *International Symposium on Computational Intelligence in Robotics and Automation*, 2003, p. 485.

[9] J. D. Johnston, "Transform Coding of Audio Signals Using Perceptual Noise Criteria," IEEE Journal on Selected Areas in Communications, Vol. 6, pp. 314-323, 1988.

[10] R. Knotts, I. Nourbakhsh and R. Morris, "NaviGates: A Benchmark for Indoor Navigation," In *Proc. of the Third International Conference on Robotics for Challenging Environments, ASCE*, 1998.

[11] D. Lee and M. Recce, "Quantitative evaluation of the exploration strategies of an intelligent vehicle," International Journal of Robotics Research, 16(4), pp. 413-447, 1994.

[12] D. Y. Pan, "Digital Audio Compression," Digital Technical Journal, 5(2): 2840, 1993.

[13] I. Rañó and J. Minguez, "Steps towards the automatic evaluation of robot obstacle avoidance algorithms," In *Proc. of Workshop of Benchmarking in Robotics, in the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2006.

[14] J. Tobias, *Foundations of Modern Auditory Theory*, Ed. New York and London: Academic Press, 1970.

[15] S. C. Wong, L. Middleton and B. A. MacDonald, "Performance metrics for robot coverage tasks," in *Proc. of the Australasian Conference on Robotics and Automation (ACRA)*, 2002.

[16] E. A. N. Carvalho, M. V. S. Alves, E. O. Freire, J. Montalvão and L. Molina, "Avaliação de Desempenho de Arquiteturas de Navegação de Robôs Móveis com Base em Padrões Humanos," In: Proc. SBA'09, 2009, paper 55826.