34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Bio-robots Automatic Navigation with Electrical Reward Stimulation

Chao Sun[1] Xinlu Zhang[2] Nenggan Zheng[3] Weidong Chen[4] Xiaoxiang Zheng[5]

Abstract— **Bio-robots that controlled by outer stimulation through brain computer interface (BCI) suffer from the dependence on realtime guidance of human operators. Current automatic navigation methods for bio-robots focus on the controlling rules to force animals to obey man-made commands, with animals' intelligence ignored. This paper proposes a new method to realize the automatic navigation for bio-robots with electrical micro-stimulation as real-time rewards. Due to the reward-seeking instinct and trial-and-error capability, bio-robot can be steered to keep walking along the right route with rewards and correct its direction spontaneously when rewards are deprived. In navigation experiments, rat-robots learn the controlling methods in short time. The results show that our method simplifies the controlling logic and realizes the automatic navigation for rat-robots successfully. Our work might have significant implication for the further development of bio-robots with hybrid intelligence.**

I. INTRODUCTION

The interaction between animals' brains and outer man-made systems through brain-computer-interface(BCI) has provided a new approach for animal robots. With intra-cortical stimulation, e.g. electrical pulse, researchers send commands to sensory cortex of animals directly. This stimulation could induce animals to finish various sorts of tasks, from simply pressing a lever to walking along complicated 3-D routes. In recent two decades, bio-robots have been realized in different kinds of creatures[1–9].

Bio-robots are prior in many aspects to traditional mechanical robots such as mobility and adaptability. This mainly benefits from the remarkable locomotory capabilities and intelligence of animals which none of man-made robots can parallel. On the other hand, because of the self-consciousness of animals, their behavioral reactions to the commands vary subject to their temporal mental and physical conditions. The controlling logic for bio-robots is difficult to model and explicate. At present, navigation for bio-robots mostly relies on the empirical guidance of human operators.

Take rats as an example. The first rat-robot was developed in 2002[3]. Researchers realized three behavioral control commands, ***Forward***, turning ***Left*** and ***Right*** to steer the rat-robot along a complicated route, even in 3-D terrains. Many rat-robots with similar principle were developed by other teams[6, 10, 11]. Meanwhile the research on the automatic navigation for rat-robots is still on early stage. The current studies focus on the expression and modeling of the controlling logic, either explicitly listing every single rule[12, 13] or implicitly building a mathematic model to learn the controlling process as a whole[14]. These methods attempt to steer rat-robots automatically with the same principles as human operators. In navigation, all three kinds of commands are involved. However, the biological mechanism of the *stimulation-reaction* process has not been fully understood, so the relationship between the intensity of stimulations and the extents of responding behaviors cannot be modeled precisely. Especially for the turning commands in rat-robots, the micro-stimulation in left or right somatosensory cortices(SI) generates 'virtual touch' in opposite side of the whisker as a cue to induce the rat-robot to turn to corresponding direction[15]. Due to this principle, the exact turning angles controlled by these turning commands could not be determined precisely. It causes that former methods for automatic navigation with the turning commands suffer from the uncertainties of the effect of these commands. On the other hand, current navigation algorithm for rat-robots controls the rat as mechanical robots with the similar methods in traditional robotics. The unique intelligence of animals, especially the remarkable spatial learning capability is ignored. It is well known that rats have remarkable ability in navigation, as piloting and dead reckoning[16]. In practice, the rat-robots have shown the capabilities that they would approach to the target for rewards and correct their behaviors when no reward is given. This intelligent behavior can be used in bio-robot navigation to integrate the intention of animals and the controlling methods into real intelligent robots.

In this paper, a new automatic navigation method for rat-robot is introduced. We apply single ***Reward*** command to steer the rat-robot, without any other turning commands (***Left*** or ***Right***). The rat-robot receives rewards timely when it performs ideal behaviors as walking along the right route. If it walks into incorrect direction, the rewards are cut off as a cue. The rat-robot would realize its error and correct to right direction in seeking for rewards. This method simplify the automatic navigation algorithm largely by taking fully advantage of the rats' learning capabilities. The results of navigation in experimental mazes shows that our method can steer the rat-robot along the target route without human interference, thereby realizes the automatic navigation for bio-robots successfully.

II. METHODS

A. Reward Seeking and Trial-and-Error in Rats

Rewards have been widely used in research on spatial learning of rodents[16]. Studies have shown that in spatial exploring

*This work was supported by grants from the National Natural Science Foundation of China (No. 61031002, 61003150), the National Basic Research Program of China (No. 2011CB504400), Zhejiang Provincial Key Science and Technology Program for International Cooperation(No. 2011C14005), the Fundamental Research Funds for the Central Universities.

[1]Chao Sun,[2]Xinlu Zhang, [3]Nenggan Zheng,[4]Weidong Chen, [5]Xiaoxiang Zheng are with Qiushi Academy for Advanced Studies (QAAS), Zhejiang University, Hangzhou 310027, China. [4]Weidong Chen is corresponding author(e-mail: qaas@zju.edu.cn).

[5]Xiaoxiang Zheng is also with the Department of Biomedical Engineering, Zhejiang University, Hangzhou 310027, China. [1]Chao Sun,[2]Xinlu Zhang are also with College of Computer Science, Zhejiang Univerisity, Hangzhou 310027, China.

experiments, rats learn to correct their behaviors if the rewards are deprived[17]. This process as trial-and-error is based on the remarkable learning ability of rodents[18, 19]. In these experiments, the rewards refer to food or water. Compare with the traditional forms of rewards, electrical stimulation is superior in the precision and realtimeness[20, 21]. In navigation experiments, the rat-robot receives the feedback immediately after the behavioral performance. In practice we found that the rat-robot can realize its incorrect behaviors if rewards are not available and return to the former location to attempt the other choices straightway. The behaviors as reward-seeking and trial-and-error could be utilized in automatic guidance for rat-robots.

B. Automatic Navigation Algorithm

We have developed automatic navigation system for rat-robots with different controlling algorithms in early work[12, 14]. In this paper, the controlling decision making module of the system adopts new automatic controlling method. Only the command ***Reward*** is involved in navigation which is introduced as follows.

Before the experiment, the panoramagram of the experimental scene is captured by a bird-eye camera. Based on this photograph, all feasible paths in this scene are extracted with the computer vision algorithms and then divided into a serial of grids. The size of each grid is determined as the average length the rat-robot progresses between the *Forward* commands in early experiments. For navigation, researchers set the target route manually or appoint the destination, with the optimal route yielded automatically by the A* algorithm which is a popular heuristic-based algorithm for path planning in robotics[22].

In experiments, the rat-robot is placed at the start point of the target route. The locomotive information of the rat is extracted from the supervisory video in real time. In this method, only the location of the head of the rat-robot is applied to support the controlling decision making, which greatly simplifies the locomotion analysis for rat-robots navigation. In each frame, the controlling decision making involves following steps:

1) We define the current location of rat-robot is $h_{x,y}$. The grid which $h_{x,y}$ within is defined as G_c. The ideal route from G_c to the destination G_d is computed by the path planning algorithm mentioned above. The G_i represents the next ideal grid along this route.
2) While the $h_{x,y}$ is still within the G_c, the rat-robot makes none movement and receives no rewards.
3) When the rat-robot enters a new gird denoted as G_n, our algorithm compares G_n with the G_i in order to check whether it walked along the right route.
4) If G_n equals G_i, the rat-robot is considered made the right decision. The reward stimulation is given to induce it to move on. The G_c is reset by the current grid G_n. Once the G_c reaches G_d, the navigation is finish as a success. Otherwise the controlling process iterates back to step 1.

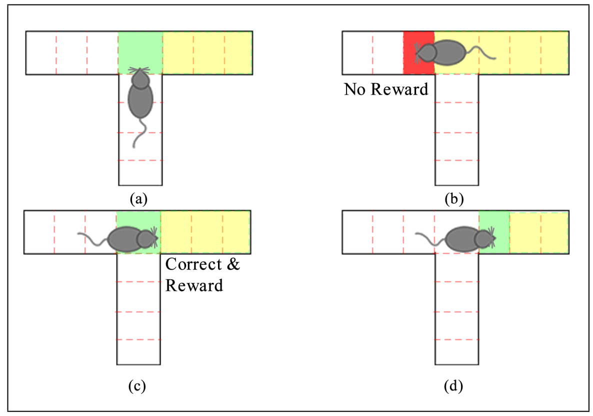


Fig. 1. An example of automatic navigation for rat-robots. We take an example in T-maze which is divided into grids. The green grids represent current correct grid the rat-robot locates and the yellow grids denote the ideal route to the target. (a) The rat-robot is placed at the bottom end of the maze. (b) When the rat-robot meets the crossing in the center of the maze, no command hints the right direction. The rat-robot walks into the arm chosen by itself. With entering the wrong grid(marked as red), no reward is given. The ideal route is re-computed. (c) Without receiving any reward, the rat-robot realizes its error, and turn around to the former grid. Because the route is reset, entering the center grid is rewarded. (d) Due to the experience of former choice and the rewards, the rat-robot try the other direction. When the rat-robot enters the right arm, rewards are given to encourage it to finish its behavior. Then the rat-robot keeps walking along the right direction to the destination.

5) If G_n and G_i are not the same grid, we define the rat-robot made a incorrect decision and no reward will be given. The G_c is set to be the current grid G_n, meanwhile the target route is computed and generated again. Note that in this situation, the ideal route may differ from the original one because the start point is reset as the current grid. This guarantees that no matter how far the rat-robot deviates from the original route, the right choice would always be rewarded to steer it turn to the destination. The controlling turns back to step 1.

A diagram of the controlling method is shown in Fig 1.

In this method, no turning commands are involved to control the turning direction for the rat-robot. In traditional navigation, the timing of the turning and rewarding commands is the major issue that influences the effective of the controlling performance. Once the rat-robot responds to the turning commands correctly, the reward should be given in time to induce it to finish this behavior. However, the exact timing differs from each rat because of their individual conditions. On the other hand, the responding behaviors should be described with many locomotive parameters, such as the positions of the head and body, the orientation of the head, the movement velocity of different parts of the body. The calculation of these parameters increases the complexity and time consumption of the controlling algorithms. In our method, this issue is successfully addressed by simplifying controlling logic and reducing the command set, leaving decisions making and correcting performed by the rat itself.

C. Subjects and Rewards

The navigation experiments are performed with adult Sprague-Dawley rats. In left and right medial fore-brain bundle(MFB) area (AP -3.8, ML +1.6, DV +8.2), one pair of bipolar electrodes is implanted. As the most frequently investigated reward-related areas, stimulation in MFB generates intensive excitement as a perfect form of rewards[23]. No electrodes for other commands are used in this method. After a 7-days post-operative recovery, the rat is placed in a operant chamber with a bar in one side. Once the bar is pressed by the rat, the MFB reward command is given immediately. Through this training, the effect of the stimulation and the rat's reward-seeking behavior is reinforced. Note that in rat-robots, only one side of MFB stimulation is chosen as the commands of ***Reward*** according to performance in the operant-chamber-tests. The difference of the side of MFB stimulation will influence the habit of rat-robot in walking. The each stimulation consists of 10 pulses with at 10ms intervals, with each pulse of 8V and lasting 1ms. More information about our rat-robot can be found in our early papers[6, 24].

III. Experiments and Results

The automatic navigation is performed in a T-maze. In early experiments, we built eight-arms maze for navigation. In this setting, five arms are blocked to establish a T-maze. The rat-robot is placed at the end of the bottom arm and supposed to be steered to either left or right arm set by human operators or the system randomly. Once the rat-robot reaches the target, this trial is labeled as success. If it walks into the end of the opposite arm, or cannot finish the navigation within 20 seconds, this trial would be defined as a failure. Because the target is chosen randomly in each trial, the experience in past trials makes no effect to the decision making and even may be misleading. The rat-robot should learn to correct immediately after wrong behaviors from trials.

4 rat-robots (No. A1-A4) are tested in T-maze automatic navigation. In each day, the rat-robot is tested in T-maze experiment paradigm with 3 groups with each consisting of 11 trials. Between the groups the rat-robot takes a two-minutes rest. The results are introduced in Table I.

TABLE I
Results of T-maze automatic navigation

Rat No	1st Day		2nd Day		3rd Day	
	CR[1]	SR[2]	CR	SR	CR	SR
A01	89.7%	90.1%	89.7%	90.1%	83.3%	93.9%
A02	21.1%	51.5%	57.9%	75.8%	63.2%	78.8%
A03	40.0%	63.6%	76.8%	84.4%	85.7%	90.1%
A04	52.9%	72.7%	75.0%	81.8%	82.4%	87.9%

[1] CR: correcting success rate. The proportion of which the rat-robot chooses a wrong direction but corrects to the ideal arm by the no-reward controlling.

[2] SR: total success rate. The proportion of which the rat-robot reaches the target in total trials, including the correcting trials and the the trials in which the initial choice is right and rewarded.

As shown in the table, our system steers the rat-robot in T-maze with great performance. The average success rate for four rat-robots reaches over 85% (87.68±6.41%) after only three days experiments. Meanwhile the average rate that rat-robots correct their incorrect behaviors is also over 75% (78.7±10.4%). The results also indicate the remarkable learning ability of rats. Some rat-robots (as A01) learns the controlling method during less than 30 trials with the success rate is over 90%. And others improve their performance in a relatively short time. The excellent intelligence of animals is fully utilized in our controlling algorithm to realize the automatic navigation tasks. Two sets of video clips of automatic navigation procedure in two days are shown in Fig 2.

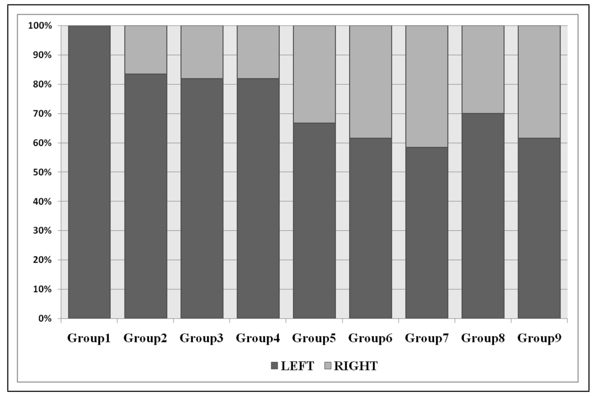
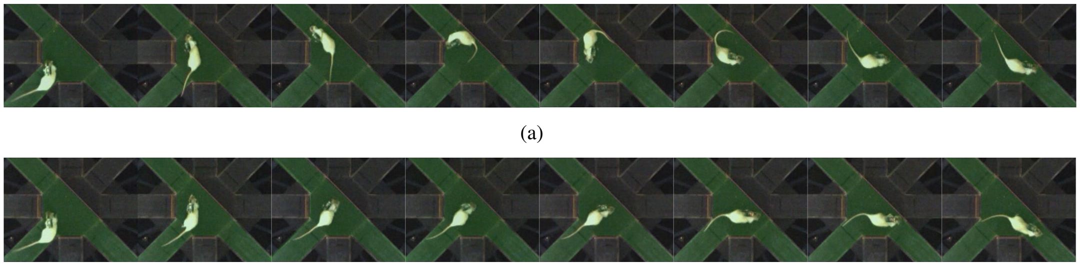


Fig. 3. The change in the preferred direction of the rat-robot (No. A03). The preference denotes which direction the rat-robot prefers to turning to when it enters a turning crossing in the maze. Note that group 1-3 refers to the first day, group 4-6 to the second day, and group 7-9 to the third day.

As mentioned before, because MFB stimulation is given in one side of the brain, the electrical stimulation influences the locomotive behaviors of the rat-robot. Normally, the rat-robot prefers to walking forward the contralateral direction of the stimulation. Fig 3 demonstrates the preferred choice in T-maze of the Rat A03 which receives the ***Reward*** stimulation in the right side MFB. At the beginning of the experiments, the rat-robot always turns to the left as the preferred choice. After navigation experiments for 3 days, the rat-robot learns the trial-and-error method in both directions. This adjustment is also implied in Fig 2. This correction of the rat-robot's preference indicates our method not only controls the temporal behaviors like making the turning choices, but also influences the long-term habits of animals.

IV. Discussion and Conclusion

This paper proposed a new controlling method for bio-robot automatic navigation with electrical reward stimulation. The complicated controlling logic was simplified by taking advantage of intelligent behaviors of animals themselves. The system gave or deprived realtime electrical rewards according to the performance of the rat-robot. Based on the reward-seeking and trial-and-error behaviors in rats, the rat-robots learned to make and adjust decisions by itself. In this way, the intelligence of animals was integrated with the controlling algorithm to realize the automatic navigation for bio-robots. This work might provide new ideas for the further studies on bio-robot controlling and intelligence hybrid.



(a)

(b)

Fig. 2. Video clips of automatic navigation (Rat A01). The green area represents the feasible paths, with other arms are blocked by baffles. Fig (a) is captured from the supervisory video of the automatic navigation at the first day. The rat-robot enters the wrong arm and then corrects to the right direction. Fig (b) is from the navigation video in the third day. The rat-robot attempts one direction without entering any arm. While no reward is given, it turns to the opposite direction. Comparing these clips in two days, it is clear that the rat-robot has learned and adapted the controlling method.

In T-maze, the automatic navigation with single reward stimulation achieves good results. In multi-choices scenes, studies show that rats would traverse all possible paths seeking for reward[16]. We also manually steer rat-robots in eight-arms maze with ***Reward*** commands and the results verify the former theory. The automatic navigation in eight-arms maze would be tested in following work.

The navigation in open field with ***Reward*** commands would be different with experimental scenes. In real-world environment, the decision making is not explicit as in turning points of any mazes. The possible solution is to design more natural route according to the instinct of rats. The rewards should be given once the rat-robot approaches forward the target. In this way, the rat-robot could be controlled to given destination in practical application.

ACKNOWLEDGMENT

The authors would thank Chaonan Yu, Fang Wang and Yan Cao for the animal surgeries and taking care of the rats.

REFERENCES

[1] R. Holzer and I. Shimoyama, "Locomotion control of a bio-robotic system via electric stimulation," in *Intelligent Robots and Systems, 1997. IROS '97., Proceedings of the 1997 IEEE/RSJ International Conference on*, vol. 3, pp. 1514 –1519 vol.3, sep 1997.

[2] Y. Kuwana, I. Shimoyama, and H. Miura, "Steering control of a mobile robot using insect antennae," in *Intelligent Robots and Systems 95. 'Human Robot Interaction and Cooperative Robots', Proceedings. 1995 IEEE/RSJ International Conference on*, vol. 2, pp. 530 –535 vol.2, aug 1995.

[3] S. K. Talwar, S. Xu, E. S. Hawley, S. A. Weiss, K. A. Moxon, and J. K. Chapin, "Behavioural neuroscience: Rat navigation guided by remote control," *Nature*, vol. 417, pp. 37–38, May 2002.

[4] S. Brown, "Stealth sharks to patrol the high seas," *New Scientist*, vol. 2541, p. 30, 2006.

[5] Y. Wang, X. Su, R. Huai, and M. Wang, "A telemetry navigation system for animal-robots," *Robot*, vol. 28(2), pp. 183–186, 2006.

[6] Z. Feng, W. Chen, X. Ye, S. Zhang, X. Zheng, J. Jiang, L. Jin, Z. Xu, C. Liu, F. Liu, J. Luo, Y. Zhuang, and X. Zheng, "A remote control training system for rat navigation in complicated environment," *Journal of Zhejiang University - Science A*, vol. 8, pp. 323–330, 2007. 10.1631/jzus.2007.A0323.

[7] W. Wang, C. Guo, J. Sun, and Z. Dai, "Locomotion elicited by electrical stimulation in the midbrain of the lizard gekko gecko," in *Intelligent Unmanned Systems: Theory and Applications* (A. Budiyono, B. Riyanto, and E. Joelianto, eds.), vol. 192 of *Studies in Computational Intelligence*, pp. 145–153, Springer Berlin, 2009.

[8] A. Bozkurt, R. Gilmour, and A. Lal, "Balloon-assisted flight of radio-controlled insect biobots," *Biomedical Engineering, IEEE Transactions on*, vol. 56, pp. 2304 –2307, sept. 2009.

[9] H. Sato, S. Kolev, N. Goehausen, M. Nyi, T. Massey, P. Abbeel, and M. Maharbiz, "Cyborg beetles: The remote radio control of insect flight," in *Sensors, 2010 IEEE*, pp. 1 –4, nov. 2010.

[10] Y. Wang, K. Yuan, J. Li, and W. Song, "A new remote-controlled animal robot training system," in *Integration Technology, 2007. ICIT '07. IEEE International Conference on*, pp. 599 –602, mar. 2007.

[11] R. Huai, J. Yang, H. Wang, and X. Su, "A new robo-animals navigation method guided by the remote control," in *Biomedical Engineering and Informatics, 2009. BMEI '09. 2nd International Conference on*, pp. 1 –4, oct. 2009.

[12] Y. Zhang, C. Sun, N. Zheng, S. Zhang, J. Lin, W. Chen, and X. Zheng, "An automatic control system for ratbot navigation," in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on Int'l Conference on Cyber, Physical and Social Computing (CPSCom)*, pp. 895 –900, dec. 2010.

[13] M. Wan, X. Huang, H. Ni, and B. Kong, "A new rat navigation method based on cc2431," in *Computer Research and Development (ICCRD), 2011 3rd International Conference on*, vol. 2, pp. 262 –265, march 2011.

[14] C. Sun, N. Zheng, X. Zhang, W. Chen, and X. Zheng, "An automatic control model for rat-robot," in *33rd Annual International IEEE EMBS Conference*, 2011.

[15] R. Doty, "Electrical stimulation of the brain in behavioral context," *Annu Rev Psychol*, vol. 20, pp. 289–320, 1969.

[16] I. Whishaw and B. Kolb, *The Behavior of the Laboratory Rat: A Handbook with Tests*. Oxford University Press, 2004.

[17] C. L. Hull and K. W. Spence, ""correction" vs. "non-correction" method of trial-and-error learning in rats.," *Journal of Comparative Psychology*, vol. 25(1), pp. 127–145, 1938.

[18] W. Thorpe, *Learning and Instinct in Animals*. Methuen, 1956.

[19] S. M. Kundey and J. D. Rowan, "Single and double alternation learning in rats: The role of set size and correction," *Learning and Motivation*, vol. 40, no. 1, pp. 1 – 14, 2009.

[20] J. Olds and P. Milner, "Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain.," *Journal of Comparative and Physiological Psychology*, vol. 47, pp. 419–427, 1954.

[21] R. A. Wise and P. P. Rompre, "Brain dopamine and reward," *Annual Review of Psychology*, vol. 40, pp. 191–225, 1989.

[22] R. Dechter and J. Pearl, "Generalized best-first search strategies and the optimality of a*," *J. ACM*, vol. 32, pp. 505–536, July 1985.

[23] L. Hermer-Vazquez, R. Hermer-Vazquez, I. Rybinnik, G. Greebel, R. Keller, S. Xu, and J. K. Chapin, "Rapid learning and flexible memory in tasks in rats trained with brain stimulation reward," *Physiology & Behavior*, vol. 84, no. 5, pp. 753 – 759, 2005.

[24] X. Ye, P. Wang, J. Liu, S. Zhang, J. Jiang, Q. Wang, W. Chen, and X. Zheng, "A portable telemetry system for brain stimulation and neuronal activity recording in freely behaving small animals," *Journal of Neuroscience Methods*, vol. 174, no. 2, pp. 186 – 193, 2008.