Hemisphere

An Intuitive Interface for Domestic Users Remote Driving a Mobile Robot

Sheng Kai Tang
Computational Design Program
Carnegie Mellon University

Abstract

How to operate remote controlled robots is always an important issue in Human Robot Interaction (HRI) field during the past twenty years. While these industrial and professional robots step into domestic environment, the interface design becomes even more important. In order to serve ordinary people instead of professional technicians, the way to operate, control and interact with domestic robots should be as intuitive as using everyday things. After reviewing existing research results, we find out that these solutions do get rid of constraints of conventional commanding interface but also create new problems. In this research, we reveal a common problem of existing interfaces. This problem is that these interfaces try to eliminate the steep learning curves by adopting natural channels of communication but create new tasks and problems for users to engage in. We seek to propose an intuitive interaction model based on Norman's affordance approach to solve this problem. In order to realize this model, we go through a seven steps methodology which mainly bases on ergonomic study and cognitive survey further extended to hardware, algorithm, and system design. The final step is a preliminary user test which leads us to assure the achievement of our goal of intuitive design and also point out problems. The most significant finding is that there are some mismatch between results of cognitive survey and hand ergonomics and could be further investigated to challenge the theory of affordance.

INTRODUCTION

Remote driving which provide a magic way of interaction between children and a toy car is not only for children's fun. In reality, remote driving is seriously applied in diverse fields. From extreme field discovery [1][8][9], battle field rescue [2][23] to industrial field assistance [24][25], we can obviously assure the important benefits of remote driving. It provides a safe way for human to virtually involve in dangerous environment and execute hazardous tasks. How to operate those remote controlled robots becomes an important issue in Human Robot Interaction (HRI) field during the past twenty years. Researchers gradually recognized that in addition to increasing the intelligence of robots, providing good ways for interacting with them is important as much [27].

While these industrial and professional robots described above step into domestic environment, the interface design becomes much more important. In order to serve ordinary people instead of professional technicians, the way to operate, control and interact with domestic robots should be as intuitive as using everyday things [13][17][18][27]. Too many mental loads such as learning or memorizing instructions are not expected. Hence, instead of text input, preprogrammed pedants and joysticks which mostly require plenty of practice to master [13], researchers attempt different expressional channels of human such as commanding by voices [11][17][18][22], planning path by sketches [4][5][6][14][21] and directing by gestures [7][3][12][16]. They assume these "natural" channels of commanding could decrease and even eliminate the learning curve.

After reviewing these research results, we find out that these solutions do get rid of constraints of conventional commanding interface but also create new problems. Voice command systems have problem relating words with similar semantics. Sketch path planning systems rely on users' spatial ability to correlate the real world and 2D maps. Hand gesture systems always force users to perform awkward actions to trigger desired functions. Theses solutions didn't take the advantage of "natural" and are not "intuitive" either. Based on Spool's idea [28], there must be some knowledge gaps in these solutions, and design could be a way to bridge gaps. From Norman's point of view [19], these solutions don't provide the "perceived affordance" for the user to reference. Hence, the research question is how to design a both natural and intuitive interface for driving domestic mobile robot remotely. The objectives are to reveal problems of existing solutions, to suggest new design strategies, and to develop a demonstrated intuitive interface as prove of concept.

PROBLEMS OF EXISTING SOLUTIONS

Verbal Commanding

The idea of verbal commanding is from the human machine interaction filed which had predicted ten years ago that conventional commanding interface in the industry would not be suitable for ordinary people to interact with future domestic robot, and the Natural Language Dialogue (NLD) could be a promising approach. Although there are still many argues about the appropriateness of NLD, the NLD proponents have made many applications and prototypes to demonstrate the possibilities of it. These applications are mostly based on the assumption that language is the most readily available means of communication and through it users can interact with robot naturally without any of the possibly steep learning curves required for conventional approaches such as GUI [17].

By surveying human linguistic behavior, Tasuno et al. [22] indicate that human always use sequence of keywords to command in our daily life such as "stretching arm," "shortening arm," "raising hand," and so on. Matsumura et al. [18] keep improving previous idea by using demonstrative pronouns combined with some simple gestures to provide more flexibility. For example, the word "hand" in the "raising hand" could be replaced by "it" which is meant to be anything the user is pointing at. Kulyukin et al. [17] adopt NDL to assist visually impaired with way-finding. They enable users to give commands with their own grammar by providing a dynamic semantic network.

Through the research results above, we can see researchers' efforts of providing a verbal interaction with robot. Enabling users to replace control targets or to command based on their own grammatical habits does gradually realize NLD. However, the problem occurring in the real situation which always results in low accuracy of recognition and misunderstand of communication is semantic ambiguity [10]. For example, "turn on" and "power up" might result in the same action only when a user and a robot share identical vocabulary knowledge, and this always requires learning and memorizing after all.

Sketch Planning

The task of commanding a mobile robot is always further translated into a path planning issue. How to give a sequence of commands which correctly leads a mobile robot to navigate in the environment and achieve goals is the common question to be solved in the field. In the past, the absolute position approach relies on predefined high accuracy map [21]. What researchers develop is to enable a robot to automatically find out the optimal path for a specific task in the provided map [20]. However, for domestic mobile robots, finding an optimal path is not important while providing novice users an easy way to real time plans a desired path which especially fits the dynamic interior arrangement of a home is very demanding [15].

Many researchers take the relative position approach for path planning because human deal with approximate relations between landmarks on the map instead of calculating accurate distances. Ferguson et al. [6] develop a sketch interface for drawing military course-of-action diagrams, which are used for strategic planning. Cohen et al. [5] propose a multimodal interface (QuickSet) enabling users to draw gestures on top of an existing map. These gestures include defining regions, specifying a route, and indicating heading. Kawamura et al. [14] provide users a way to specify a robot path by selecting intermediate points on a sketch of the environment. Chronis and Skubic [4] propose a sketch based navigation system which enables users to sketch a route map to direct the robot. Novice users only need to draw landmarks and the desire route, and the system will interpret the sketched map into spatial relations.

The sketch path planning approach is indeed a natural way for human describing their environment and planning way finding strategies. This approach based on the assumption that users are capable of sketching accurate relations of environmental landmarks. However, as we all know, besides those who are born with good spatial abilities or are professionally trained, many people have trouble to translate a 3D visual scene into 2D spatial relations. Even though for those who are fine with spatial ability, some instructional exercises are still required. This natural approach, again, eliminates the learning of using conventional control devices but create a new translation task for users to engage in.

Gesture Control

The research result indicated that to transfer our own will, we use the verbal communication channels at the 35% ratio, while the non-verbal communication channels are at the 65% ratio [18]. Among these 65% non-verbal communication channels, gesture is the most compact mean of relaying geometric information [16]. Hence, many of researchers in the HRI field take this approach to improve human robot interaction.

Fong et al. [7] present a virtual joystick technique to drive remote vehicle and this technique is to recognize user’s hand position and size. Cerlinca et al. [3] dynamically allocate the hot regions for commanding by detecting a user’s head. This approach provides free movements for the user within the camera scene instead of sticking at a rigid spot. Hu et al. [12] develop a system to recognize different patterns of gestures instead of only detecting hot zones. This approach is further extended for the development of recognition of sign language. Kortencamp et al. [16] propose a solution for recognizing spatial gestures. His approach is to detect joins of an arm and reconstruct a whole arm in 3D virtual world.

The common problem of results described above is performing awkward and hard to remember postures when commanding. Even though for the easier case which just requires users to trigger the right regions for desired commands, memorizing is still needed. The more complicated gestures are designed, the more mental loads are needed when operating. This, again, adopts another natural channel but still requires extra learning and adaptation. Although not being a novice user, problems of ambiguity still exist.

STRATEGIES FOR INTUITIVE INTERFACE DESIGN

After surveying existing solution for remote controlling a mobile robot, we come to a conclusion that the approach of natural communication channels doesn't guarantee against extra learning and adaptation. Based on our point of view, there is nothing wrong with the approach but just something missing in the design processes. In other words, solutions described above all focus on dealing with technical issues instead of proposing a suitable interaction model. That makes creating new problems is inevitable. Compensating for new problems definitely requires users to adapt and learn. In the following sections, we are going to propose some strategies for designing our intuitive interface for remote controlling a mobile robot.

Defining An Intuitive Interaction Model

According to Spool's idea [29], what make an intuitive interface are two conditions. First, the users know everything they need to operate a design and complete their objective when walking up to it. Second, the users have known ideas of how to operate a design but they are completely unaware the design is helping them bridge the gap.

We base on the second condition to propose our ideal model for an intuitive interaction design which is a two-way communication between a user and an interface. At the first step, a user can acquire or be informed of operational clues from the interface unawares. Second, a user can give commands through natural communication channels. Third, once a command is given to the interface, users can receive prompt feedbacks to assure their actions (Figure 1).

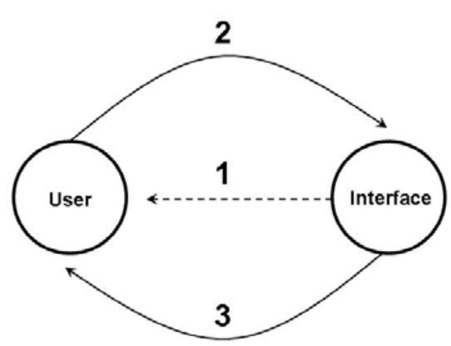


Figure 1. A model of intuitive interaction

Providing Clues

Ambiguity is a common problem existing in solutions described above especially in voice and gesture command systems. Being difficult to know the shared vocabularies and gestures results in problems, and we try to seek a way to solve it. The idea we propose for disambiguating is providing clues. Such as reminder for the password system, it successfully reminds users their password by providing related clues.

In order to achieve the goal of intuitive, we provide unaware clues for users based on Norman's idea of "perceived affordance. [19]" Norman believes that the form of a design with perceived affordances is to provide clues of functional operations. These clues consist of physical and material constraints of a form. These clues mentally inform users their abilities of manipulating everyday objects unawares. Hence, providing a physical or visible shape containing operational clues is very crucial for our intuitive design.

Adopting Natural Communication

Using natural channels for communication has benefits especially in the aspect of eliminating learning cost. Gesture is the most visible and indicative way among the 65% non-verbal communication. Comparing to verbal communication, gesture also has less ambiguity problems. Hence, for remote controlling a mobile robot, we decide to adopt gesture as our approach.

In our gesture solution, we will prevent from awkward and unnatural design of gestures such as performing special patterns or acting unnecessary big scale postures. There is even no need to memorize gestures, because ideally those gestures in our system have already been mentally existed in users' minds.

Giving Feedbacks

It is a very common scenario in our daily life, especially in those commanding solutions previously described, that people are frustrating with an interface, commanding the same action again and again, and waiting for something unknown. It could be that the system can't recognize the ambiguous commands, but it could also be that users are not sure what's wrong with their commands and wait for feedbacks.

Feedback is indeed an important element among interactions and communications. Human waits for feedbacks as references for the next action in our everyday conversations. Feedbacks also provide clues for human to assure their completeness of actions. Without anticipated feedbacks, the interaction or communication will stop and turn into a "debugging" mode. In our design, we will fulfill both functions of feedback.

METHODOLOGY AND STEPS

Based on the model and strategies proposed above, we have seven steps to realize an intuitive interface. In step one, we propose a conceptual model for our intuitive interface for remote driving which is based on the intuitive interaction model. In step two, we have an ergonomic study of hands to look for suitable geometry as our design primitive. In step three, we conduct a cognitive survey to seek for the "affordances" of this chosen primitive. In step four, these affordances are further turned into our action design. In step five, we follow the strategies and embed proximity sensors and LEDs in the chosen primitive to create the hardware. In step six, we propose and implement pattern recognition algorithms for recognizing commanding movements. In part seven, we invite three users to test our interface. This test doesn't cover all subjects of usability but focus on the issue of "intuitive".

Conceptual Model

We propose a conceptual model, or we can say a design model, of our intuitive interface for remote driving. There are two purpose of this model in this research. First, the same as other projects, it is to envision an ideal workflow of our system including proper inputs and desired outputs [19]. Second, it will be physically prototyped (system image), tested by users and compared with users' mental models. This comparison will be further used to discover the issue of "intuitive."

In the first step, user's eyes perceive both "perceived affordances" of the interface and the control target (a mobile robot). In the meanwhile, the user generates ideas of how to control the target through the interface in a very short interval. In the second step, the user performs movements on the interface by hand. In the third step, the interface gives prompt feedbacks to the user. In the fourth step, the interface recognizes the user's movement and triggers the related action of the target. In the fifth step, the target performs the action as another feedback for the user (Figure 2).

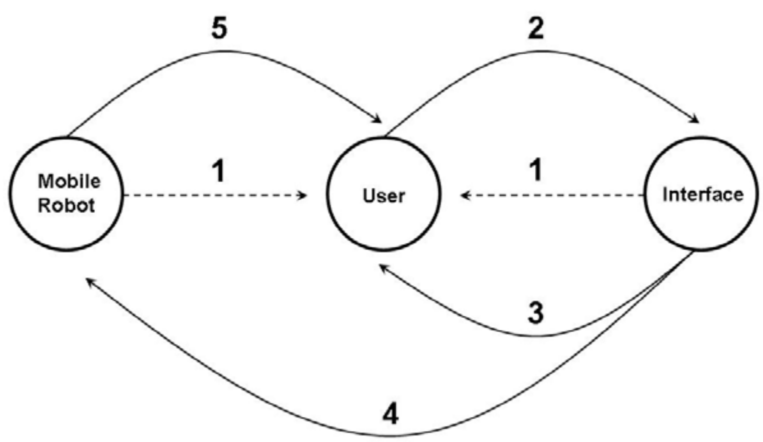


Figure 2. A model of intuitive interface for remote driving

Hand Ergonomics

Based on our strategies, our first goal is to design an interface which has a physical form providing affordances as operational clues. Second, we want to adopt natural gestures which require no 'manipulation' of an interface. The way we propose for these two conflict goals is a touch free interface. Without touching, users can perform free gestures above the interface which visually provides clues.

Ergonomics is a very important issue when developing products. It aims at providing a safe, healthy and comfortable working status. Efficiency will also increase as a result of ergonomic studies before developing and implementing a real product. In this project, we are going to choose our target geometrical primitive based on our ergonomic study of hands.

After observing and measuring 10 females and 10 males whose ages are from 25 to 35 years old, we have two ergonomic findings. First, subjects' hands are not stretched-strait and they naturally form a dorm shape most of the time. Second, the average diameter of the "hand dorm" is from 10 cm to 14 cm for our subjects (Figure 3). Based on above, we plan to adopt a dorm shape with 12 cm diameter as our primitive to design our interface. We call this dorm "Hemisphere."

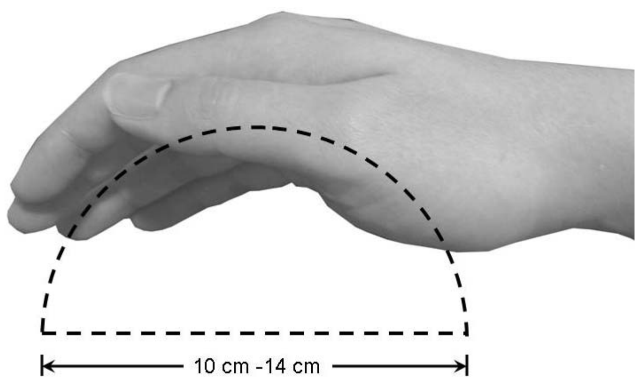


Figure 3. A result of a ergonomic study of hand

Cognitive Survey

In this step, we invite 20 subjects and conduct a survey which provides three images of our chosen primitive including perspective, top and side views of it. Based on these images, we ask subjects to draw down their intuitive solutions for four different commands which are forward, backward, turn left and turn right. These commands are basic actions for remote driving a car.

For move forward and backward, there are 56% of subjects performing path A which starts from the top (center) of the dorm to the edge of it. 44% of subjects perform path B which moves from the edge through the top of the dorm then to the opposite edge (Figure 4).

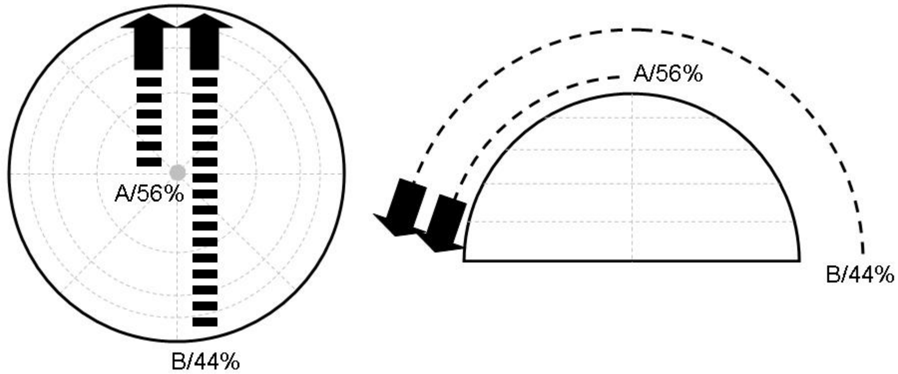


Figure 4. A statistical result of hand movements of move forward and backward

For turn left and right, there are 28% of subjects performing path A which is the same as the previous path A, and 44% of subjects performing path B which is also the same as the previous path B. However, 28% of subjects performs path C which takes the top of the dorm as center point and rotate their hands alone the edge of the dorm (Figure 5).

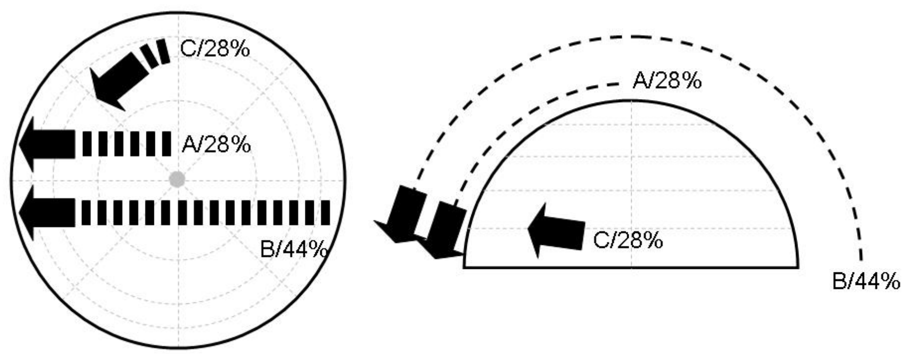


Figure 5. A statistical result of hand movements of turn left and right

Based on this survey, we are not going to figure out the only dominate affordance of the primitive but look for all affordances which might possibly existing in this primitive. We anticipate this strategy can solve the ambiguity problem.

Action Design

With the result of the previous step, we are able to plan suitable and intuitive interactions between users and this interface. Our strategy is to include all the affordances we found through the survey, therefore, this interface can afford users who might perceive an identical interface differently.

The detail planed interactions are as below. A, B and C are three cases for four basic actions which are move forward, move backward, turn left and turn right. In case A, each action starts from the top of the dorm and move to the edge directionally. In case B, each action starts from the edge and move to the opposite edge of the dorm. In case C, move forward and backward are from the top of the dorm to the edge while turn left and right take the top of the dorm as center and rotate along the edge (Figure 6).

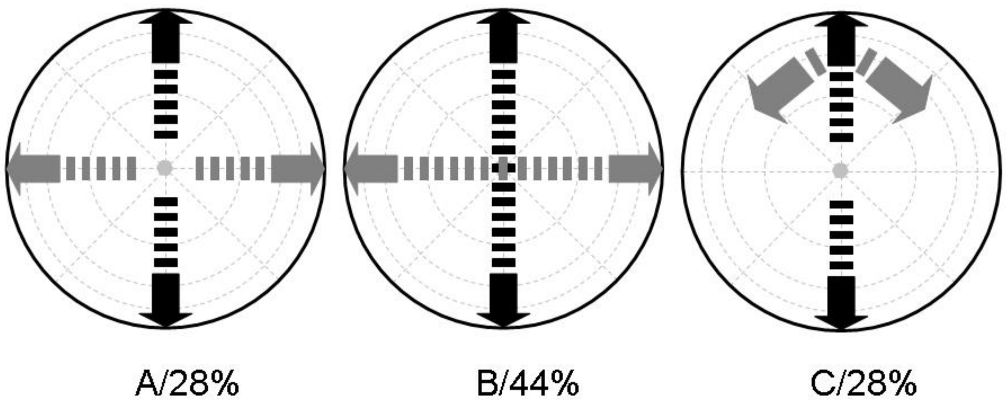


Figure 6. Different cases of action design

Hardware Design

In order to achieve the goal of touch free, we adopt the infrared (IR) sensor to sense hand movements. An IR sensor is a proximity sensor consisting of a pair of transceiver and receiver. The transceiver emits directional infrared light to the environment and the receiver receives infrared light reflected from environmental objects. The distance between the IR sensor and environmental objects will result in the intensity of reflected infrared light received by the receiver. Meanwhile, the intensity will turn into measurable voltages or currents read by a microcontroller.

We also design a module which consists of an LED and an IR sensor to provide both sensing and feedback. When a transceiver is power up (1), it emits infrared light to the environment (2). This infrared light will reflect and be received by the receiver when encountering a target (3). The voltage change is measured by a microcontroller (4). The microcontroller sends a voltage pulse to light up an LED (5). Finally, the LED will be perceived by the user as a feedback (6) (Figure 7).

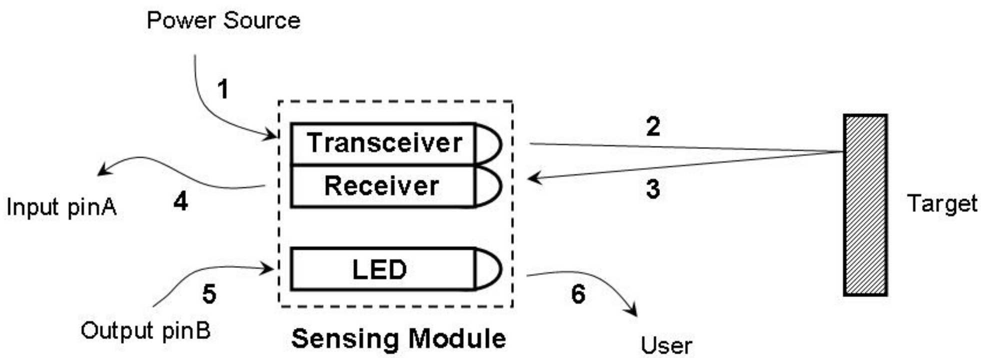


Figure 7. A design of the sensing module

We then deploy this sensing module based on the three modes proposed in the previous section. For mode A and B, we deploy 8 sensors in a cross shape on the dorm to sense directional hand movements. For mode C, we deploy another 8 sensors in a cross shape rotated 45 degrees from the previous cross to sense the hand rotational movements (Figure 8, 9).

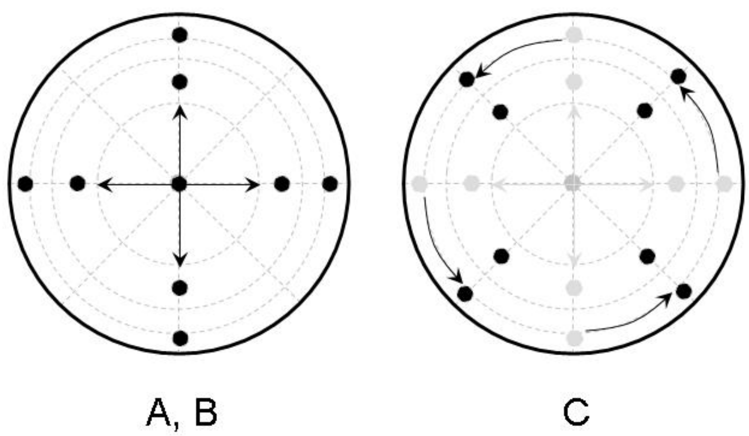


Figure 8. A deployment method of sensing modules

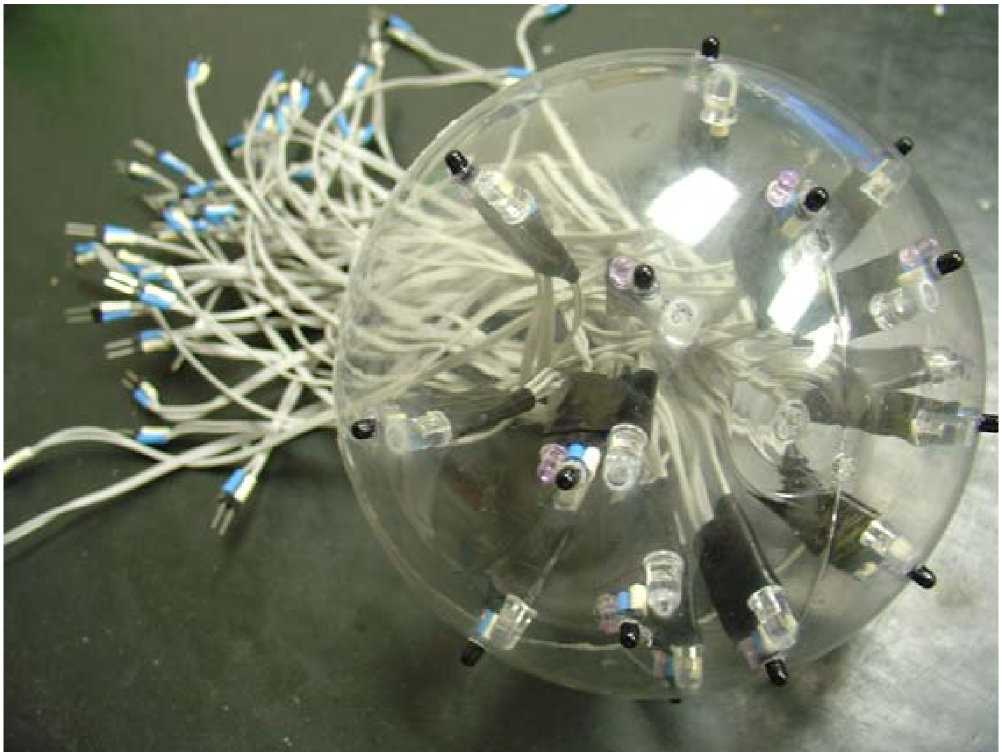


Figure 9. The result of the sensor deployment

For circuit design, each module with a transceiver, a receiver and an LED has 2 leads to the ground, 2 leads to the 5V power source, 1 lead to input port and 1 lead to the output port of a microcontroller (Figure 10). For 16 modules, we will need 16 input pins and 16 output pins of the microcontroller to read in and send out signals.

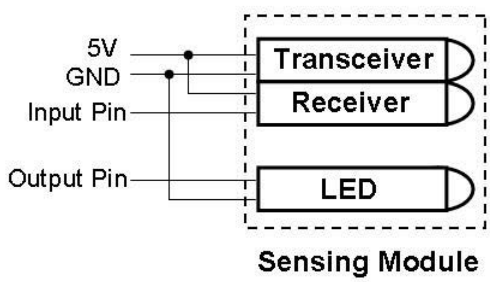


Figure 10. The circuit of sensing module

However, on one hand, the microcontroller we use (ATMEGA168) doesn't have that much digital input pins, and on the other, using out all possible pins for 16 identical modules is not efficient. Hence, we adopt the technique of Row-Column Scanning (ROS) to reduce the required amount of pins (Figure 11). The mechanism is that when powering up each row, the microcontroller scans every column (4 columns) one time to acquire data. Once every row is powered up in turn, the microcontroller will get all 16 data values.

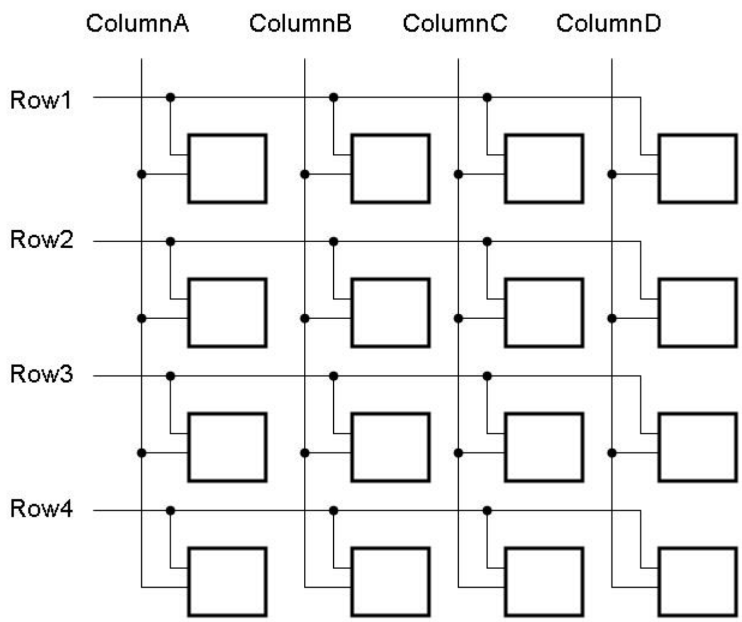


Figure 11. The Row Column Scanning technique

In order to realize ROS, we modify our circuit of each single module as shown in Figure 12, and arrange our 16 modules as a 4 by 4 matrix. By doing this, we only need 8 output pins and 8 input pins for triggering LEDs and acquiring data from receivers in 16 modules. The final result of circuit design is shown in figure 13.

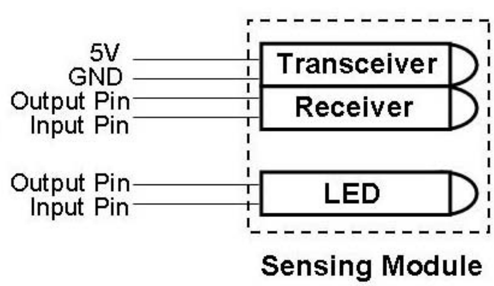


Figure 12. A modified circuit of sensing module

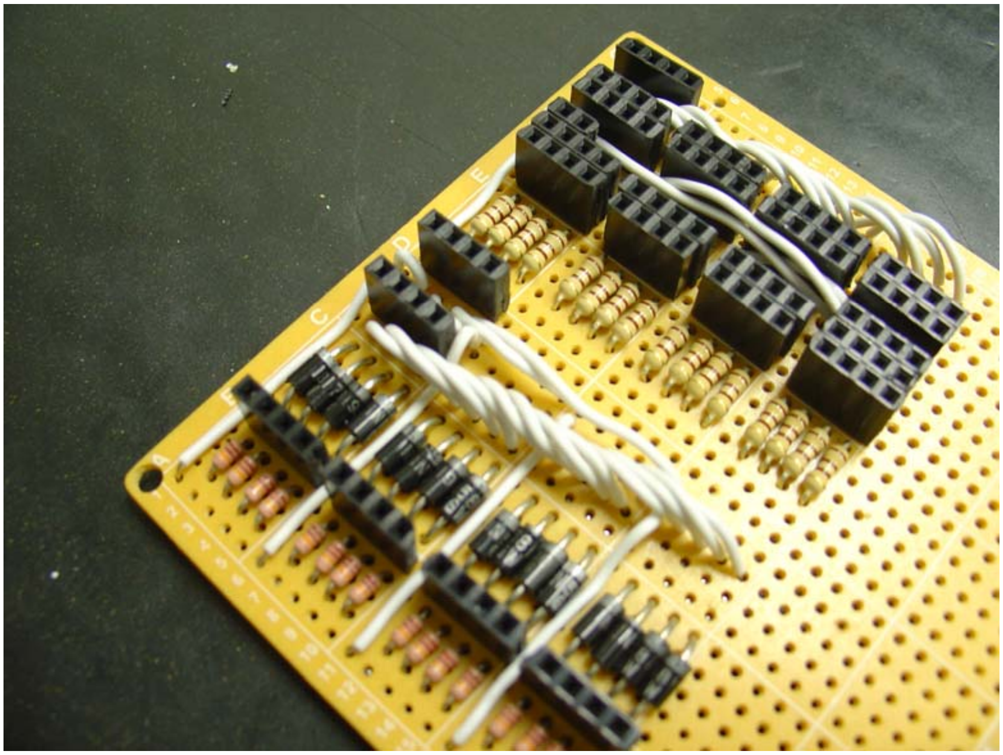


Figure 13. The result of the ROS circuit

Algorithm Design
With our hardware design, we are able to collect digital signals representing user's hand movements. However, how to design an algorithm to translate those collected digital signals is the main issue in this section. We adopt two approaches to design our algorithms. One is Pattern Matching (PM), the other is Direction Recognition (DR).

Before designing algorithms, we have to deal with our input signal data. Actually, each signal value from a sensor is from 0-255. 0 means there is no distance between a sensor and an object. 255 means the object is too far to be sensed. There is a liner relation between the value and the real distance. Based on our testing, the value is about 100 when the distance is around 3 cm. Hence, we set a threshold of 120 to convert the analog data into binary value. If the value is lower than 120, it is converted 1. If the value is higher than 120, it is 0.

For PM approach, we pre-defined all signal data of potential movements. A movement consists of a string of 80 binary values, because we use 5 sequential sets of 16 binary signals to represent a movement (Figure 14). Every time when a new 5 sequential sets of data is collected, it is compared with all pre-defined data. The pre-defined data with highest scores will be predicted as the match movement (Figure 15).

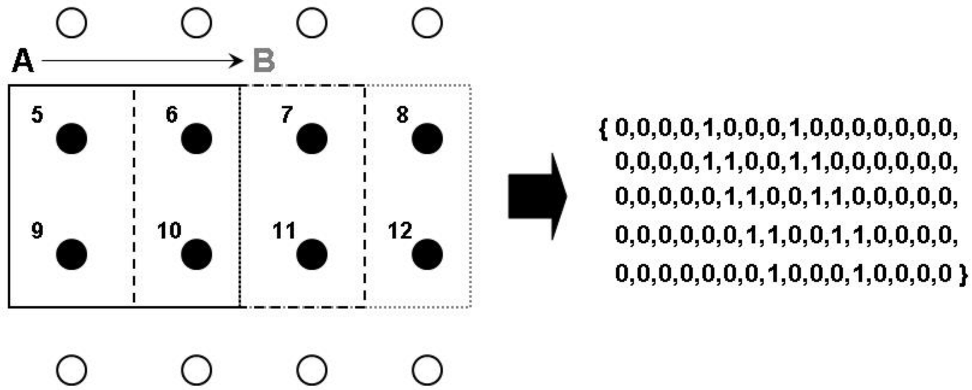


Figure 14. A concept of the PM approach

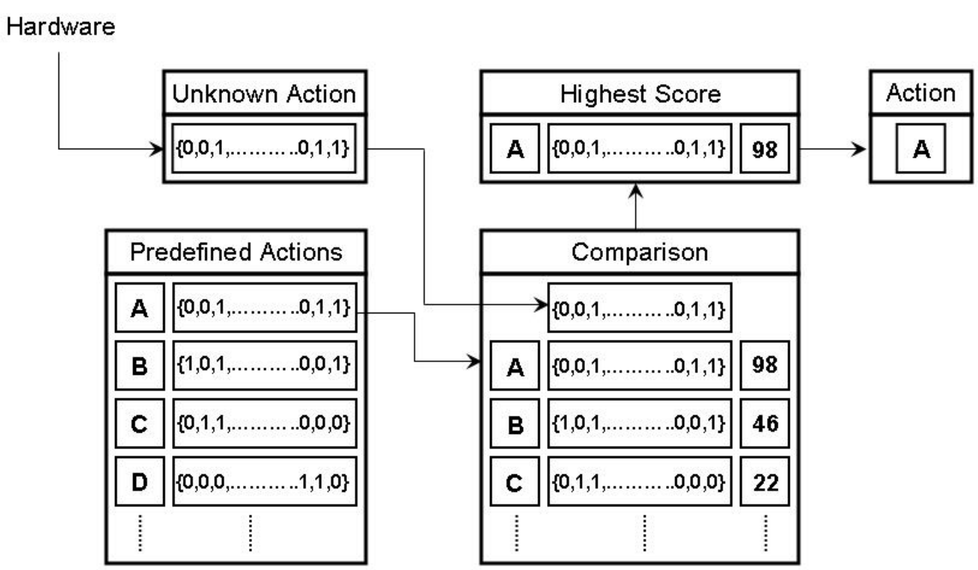


Figure 15. The detail data flow of PM approach

For DR approach, instead of predefining any data in the memory, the system always keeps two data sets which are the "now" and the "previous". Each data set consists of 16 binary values and each value represents the status of a sensor node. Each node has 4 adjacent nodes which are up, down, left and right. Every cycle, we overlap two sets of data and calculate adjacent status of each triggered nodes in 'previous' set. The adjacent status with highest amount is the actual movement (Figure 16).

For example, A and B are consisted of an unknown movement. We firstly overlap triggered nodes of both A and B. Then, we calculate the occupational condition of four adjacent nodes of each triggered nodes in A. We figure out that there are 4 rights and 2 ups. Hence, the direction from A to B is “right” based on the calculation result and it fits the actual movement (Figure 17).

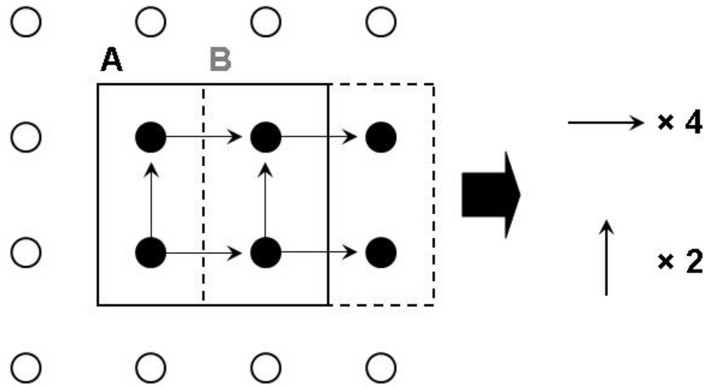


Figure 16. A concept of the DR approach

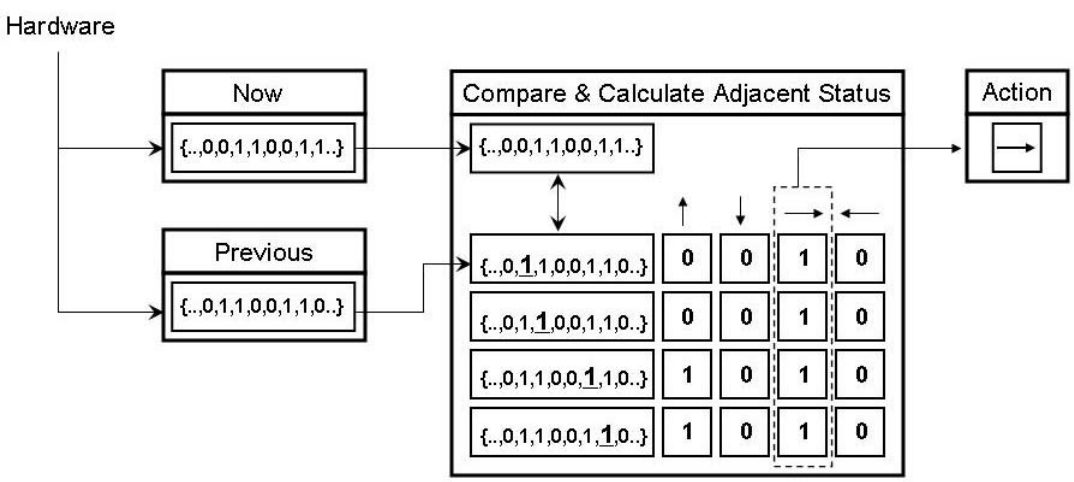


Figure 17. The detail data flow of DR approach

System Design

In our hardware system, we have the designed interface called ‘Hemisphere’, an Atmega168 microcontroller, a pair of XBee wireless modules, and a Firefly Mobile Robot Kit. The Hemisphere senses a user’s hand movements and sends sensor signals to Atmega168 microcontroller. After computing and predicting by our recognition algorithm in the microcontroller, the microcontroller sends out the predicted action to the mobile robot through XBee wireless modules. Once the robot receives the action commands, it will perform related actions (Figure 18).

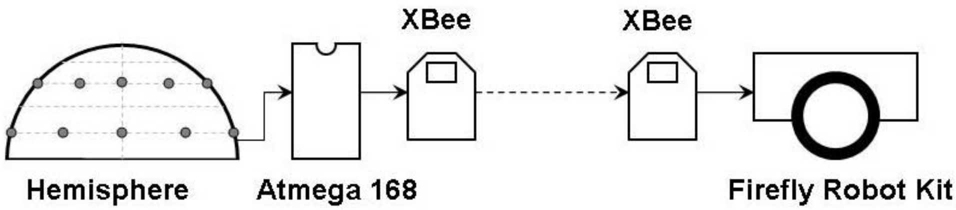
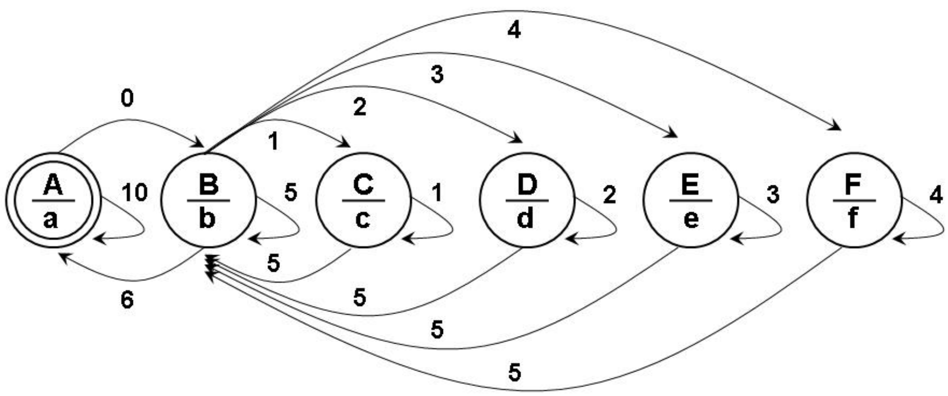


Figure 18. A system design of hardware of Hemisphere

In our application system, we implement a finite state machine combined with our proposed algorithm. In this state machine, there are 6 states, 6 actions, 11 transitions and 7 transition conditions as listed below. For example, the state is A (no coverage) when nothing is detected by the interface and the action a is no action. When something is detected by the interface, the transition condition 0 is fulfilled; the state will transit from A to B. B (no movement) state means there is no movement recognized but something is detected. If a move up is recognized, the transition condition 1 is fulfilled; the state will transit from B to C. C (up) state means a move forward gesture is detected and the action c (move forward) will be sent out (Figure 19).



States	Actions	Transition Conditions
A: no action detected	**a:** -	**0:** coverage detected
B: no movement	**b:** -	**1:** up detected
C: up	**c:** move forward	**2:** down detected
D: down	**d:** move backward	**3:** left detected
E: left	**e:** turn left	**4:** right detected
F: right	**f:** turn right	**5:** remain the same
		6: no coverage

Figure 19. A system design of software of Hemisphere

User Test
After implementing the Hemisphere, we conduct a preliminary user test which invites three subjects to perform tasks. The goal of this test is to explore the issue of "intuitive" instead of all related usability questions. We inform subjects of two things: first, this interface is touch free, and all they need to do is to hover on it. Second, please conduct actions one by one instead of sequential actions (Figure 20).

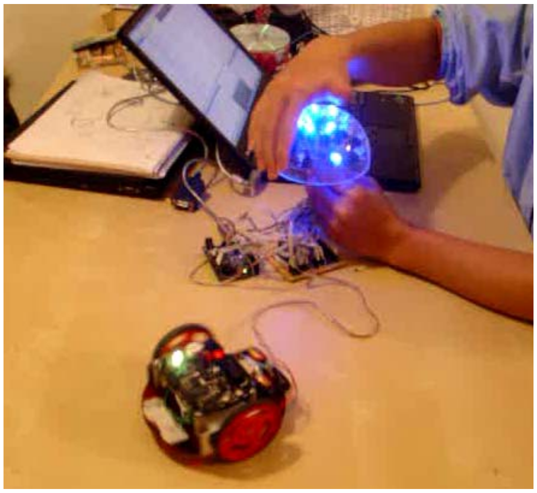


Figure 20. The user test

The tasks are 'move forward,' 'move backward,' 'turn left,' and 'turn right. Then we ask them to perform five times for each action, and calculate the amount of success and failure. The results are shown in figure 21.

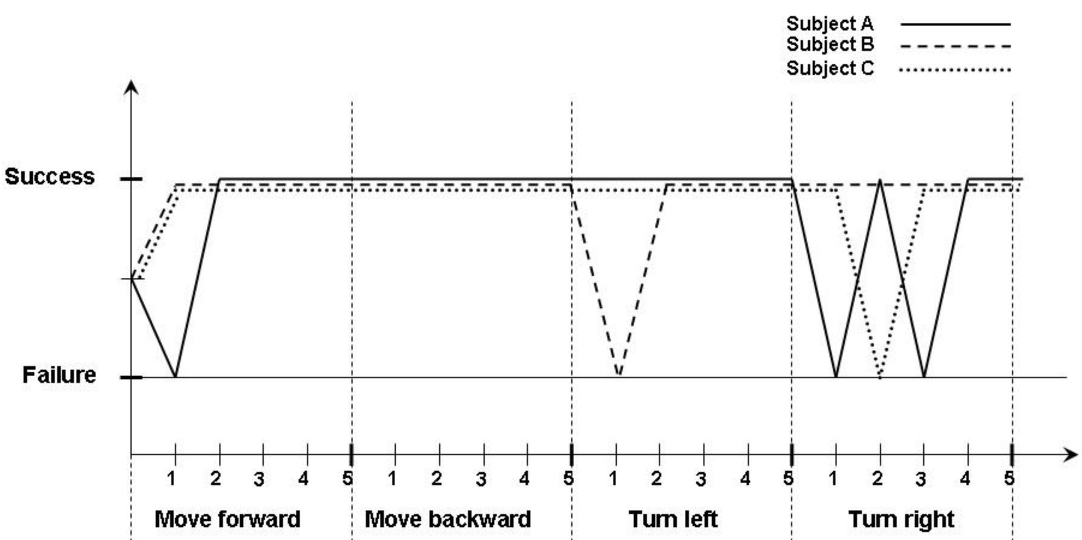


Figure 21. The result of user test

DISCUSSION

Based on our limited samples of user tests, we try to figure out some phenomena and issues of our Hemisphere interface and also discuss whether we achieve the goal of 'intuitive' design.

Intuitive

In table 1, we can see the static result of accuracy. Actually, without teaching our subjects anything about commanding, we can say our interface is quite intuitive due to the average accuracy of each action is higher than 80%. Take subject B and subject C for example, when the first time encountering this interface, they have no problem performing successful 'move forward' actions; although Subject A requires one time to try and error.

	Subject A	Subject B	Subject C	Average
Move forward	80%	100%	100%	93%
Move backward	100%	100%	100%	100%
Turn left	100%	80%	100%	93%
Turn right	67%	100%	80%	83%

Table 1. The statistic result

Affordance

Our expected actions design is based on a cognitive survey which mainly looks for "affordance" of this dorm shape. After observing our three subjects, we find out that subject A and subject C perform Case C while subject B performs Case A. Although subject A and C's actions are slightly different, their actions still fit general trend of our design. Case B is not presented in this survey might be due to the limited amount of subjects (Figure 22).

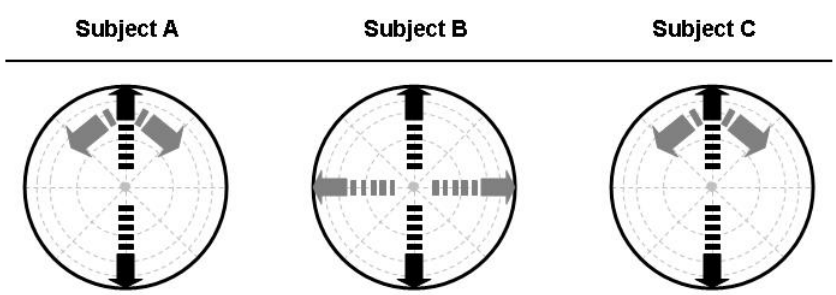


Figure 22. The result of action design case

Ergonomics
When looking into these low accuracy cases, we find out some reasons for them. It might not be the affordance problem but ergonomics ones. If the low accuracy is due to unfamiliarity, there must be a drop every time when a new actions start. However, the drop of accuracy doesn't follow this assumption at all. Neglecting move forward which three subjects act the same path for, we can easily see that the rest of low-accuracy actions relates to the paths design (Figure 23). After retesting by the author, we figure out that there are indeed some ergonomic problems resulting in low accuracy of movement recognition of the algorithm.

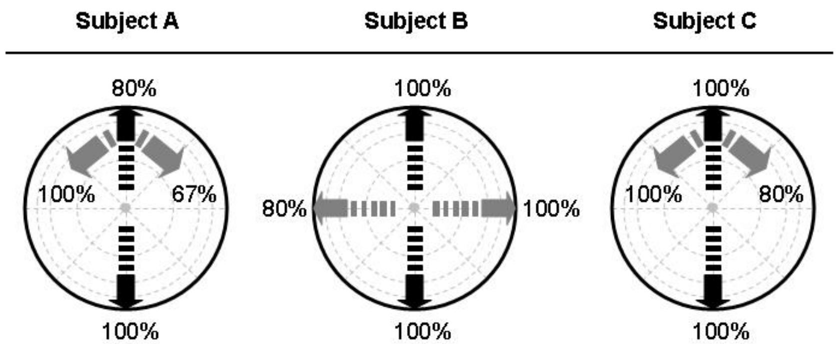


Figure 23. A fusion of statistic results and action cases

CONCLUSION
In this research, we reveal a common problem of existing interfaces. This problem is that these interfaces try to eliminate the steep learning curves by adopting natural channels of communication but create new tasks and problems for users to engage in. We seek to propose an intuitive interaction model based on Norman's affordance approach.

In order to realize this model, we go through a seven steps methodology. This method mainly bases on ergonomic study and cognitive survey further extended to hardware, algorithm, and system design. The final step is a preliminary user test which leads us to assure the achievement of our goal of intuitive design and also point out problems.

The most significant one is that there are some mismatch between results of cognitive survey and hand ergonomics. The reason for low accuracy of some actions might be because that affordance a shape provides might not fit the ergonomic performance of human. This finding could be further investigated to challenge the theory of affordance.

We only have three samples of user test and this might result in the accuracy and reliability doubts of our study. Acquiring more testing samples to increase reliability of our research will be the future study.

Reference

1. Bresina, J., Jónsson, A., Morris, P., and Rajan, K. Mixed-Initiative Planning in MAPGEN: Capabilities and Shortcomings. International Conference on Automated Planning and Scheduling, 2005.
2. Battlefield Robots: to Iraq, and Beyond, http://www.defenseindustrydaily.com/battlefield-robots-to-iraq-and-beyond-0727/
3. Cerlinca, T.L., et al. Hand posture recognition for human-robot interaction. Proc. of the 2007 workshop on Multimodal interfaces in semantic interaction, 2007, 47-50.
4. Chronis, G. and Skubic, M. Sketch-Based Navigation for Mobile Robots. Proc. of the IEEE 2003 International Conference on Fuzzy Systems 2003.
5. Cohen, P.R., et al. QuickSet: Multimodal Interaction for Distributed Applications. Proc. of the Fifth International Multimedia Conference, ACM Press (1997), 31-40.
6. Ferguson, R., Rasch, R., Turmel, W., and Forbus, K. Qualitative Spatial Interpretation of Course-of-Action Diagrams. Proc. of the 14th International Workshop on Qualitative Reasoning, 2000, .
7. Fong, T., Conti, F., Grange, S., and Baur, C. Novel interfaces for remote driving: gesture, haptic and PDA. SPIE Telemanipulator and Telepresence Technologies VII. 2000.
8. Fong, T., Bualat, M., Edwards, L., Flueckiger, L., Kunz, C., Lee, S. Y., Park, E., To, V.,Utz, H., Ackner, N., Armstrong-Crews, N., and Gannon, J. Human-Robot Site Survey and Sampling for Space Exploration. Proc. AIAA Space, 2006.
9. Fong, T., Deans, M., Lee, P., and Bualat, M., Simulated Lunar Robotic Survey at Terrestrial Analog Sites. Proc. 38th Lunar and Planetary Science Conference, 2007, Abstract 1487.
10. Frankie, J., Rayner, M., and Hockey, B. Accuracy, coverage, and speed: What do they mean to users. Proceedings of the Workshop on Natural Language Interfaces. 2000.
11. Fransen, B., et al. Using vision, acoustics, and natural language for disambiguation. Proc. HRI 2007, ACM Press (2007), 73-80.
12. Hu, C., et al. Visual Gesture Recognition for Human-Mahicne Interface of Robot Teleoperation. Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2003. 1560-1565.
13. Iba, S., Paredis, C.J.J., and Khosla, P.K. Interactive multimodal robot programming. The international journal of robotics research 24, 83 (2005), 83-104.
14. Kawamura, K., Koku, A.B., Wilkes, D.M., Peters, R.A. II and Sekmen, A. Toward Perception-Based Navigation Using Egospheres. International Journal of Robotics and Automation (2002).
15. Kim, H., et al. User-centered approach to path planning of cleaning robots: analyzing user's cleaning behavior. Proceeding of the ACM/IEEE international conference on Human-robot interaction. 2007.
16. Kortenkamp, D., Huber, E., and Bonasso, R. P. Recognizing and interpreting gestures on a mobile robot. National Conference on Artificial Intelligence. 1996, 915-921.
17. Kulyukin, V.A. On natural language dialogue with assistive robots. Proc. HRI 2006, ACM Press (2006), 164-171.
18. Matsumura, I., Tatsuno, J., Kokubo, Y., and Kobayashi, H. A man-machine interface with verbal communication: command with demonstrative pronoun. Proc. Robot and Human Communication, 1997, 398-401.
19. Norman, D. The Psychology of Everyday Things. Basic Books, 1988.
20. Schmidt, G. and Hofner, C. An Advanced Planning and Navigation Approach for Autonomous Cleaning Robot Operations. Proceedings of the 1998 IEEE/RSJ International Conference of Intelligent Robots and Systems. 1998.
21. Skubic, M., Bailey, C., and Chronis, G. A Sketch Interface for Mobile Robots. Proc. of the IEEE 2003 Conf. on SMC. 2003, 918-921.
22. Tatsuno, J., Matsuyama, S., Kokubo, Y., Kobayashi, H. Human Friendly Teaching for Industrial Robot. Proc. of 5th IEEE International Workshop on Robot and Human Communication (1996), 456-460.

23. The Bear Robot to rescue, http://blogs.zdnet.com/emergingtech/?p=338
24. Todd, D.J. Fundamentals of Robot Technology. John Wileyand Sons, 1986, chapters 3, 7.
25. Voyles, R.M. and Khosla, P.K.. Tactile Gestures for Human/ Robot Interaction. Proc. of IEEE/RSJ Conf. on Intelligent Robots and Systems, 1995, 7-13.
26. Voyles, R.M. and Khosla, P.K.. Multi-Agent Gesture Interpretation for Robotic Cable Harnessing. Proc. of IEEE Conf. on Sys., Man, and Cybe, 1995, 1113-1118.
27. Waldherr, S., Romero, R., and Thrun, S. A gesture based interface for human–robot interaction. Autonomous Robots 9,2 (2000), 151-173.
28. What Makes a Design Seem 'Intuitive'? http://www.uie.com/articles/design_intuitive/.