ComManTel 2013 Sponsors & Supporters

The Executive and Organizing Committees of ComManTel 2013 Conference greatly appreciate the Financial Sponsors and the IEEE technical Sponsorship as well as the technical sponsorship of IEEE Vietnam Section

---

Technical Sponsors: IEEE and IEEE Vietnam Section

IEEE

IEEE Vietnam Section

Co-Organizer and Financial Sponsor

Duy Tan University, Vietnam

http://www.dtu.edu.vn/default.aspx?lang=EN

Financial Sponsor

The National Foundation for Science and Technology Development (NAFOSTED), Vietnam

NAFOSTED National Foundation for Science & Technology Development

http://www.nafosted.gov.vn/en/

Co-Organizer: REV/REH

Message from the ComManTel 2013 Chairs

On behalf of the Technical Program Committee, we welcome all of you to the IEEE Computing, Management and Telecommunications Conference (IEEE ComManTel 2013) in the beautiful REX Hotel, Ho Chi Minh City, Vietnam! We are indeed delighted that the conference aims to providing a premier forum for presentation of research results and experience reporting on the cutting edge research in the general areas of computing, telecommunications, management and their applications. This year, we received around 120 submissions from more than 20 countries. Each paper received at least three peer technical reviews, comprised of more than 120 TPC members from academia, government laboratories, and industries. After carefully examining all the received review reports, the IEEE ComManTel 2013 TPC finally selected only very high quality papers for presentation at the conference and publication in the IEEE ComManTel 2013 proceedings. The first day will start with a keynote speaker chosen from renowned world-class leaders in the area. On Monday, Prof. Mounir Hamdi, IEEE Fellow and Head of the School of Computer Science and Engineering in Hong Kong University of Science and Technology, Hong Kong, China will talk about the "Massive Data Centers for Future Cloud Computing Applications." Right after the Keynote, Prof. Markus Fiedler from the Blekinge Institute of Technology, Sweden, will deliver a free Tutorial to all attendees on "Quality Economy and Sustainability of Network Services." This year, the technical sessions reflect the growing interest in a wide range of spectrum, including wireless communications and networks, computing aspects, management and applications of these technologies in our daily lives. Outstanding papers will be selected and invited for Special Issues in well-known international journals. Our objective in the future is to increase the quality and therefore reduce the acceptance rate further to reach around 30%. In addition, we would like to increase the number of Tracks to cover other areas and introduce Workshops that of interest to meet the conference theme.

A conference of this size cannot be organized without the hard work and dedication of many people. We want to thank all authors who submitted their manuscripts, as well as all the Track Chairs/Co-Chairs, TPC members, invited reviewers for all their contribution to the conference. Special gratitude goes to the IEEE Vietnam section and the IEEE headquarters for their technical sponsorship. We would also like to recognize our Co-Organizer and Financial Sponsor, **Duy Tan University** in Vietnam. Many thanks also goes to our financial sponsor: The National Foundation for Science and Technology Development (**NAFOSTED**) in Vietnam as well our Co-Organizer, **REV/REH** in Vietnam.

We hope you all enjoy the city of Ho Chi Minh and the IEEE ComManTel 2013 technical program and find this conference a productive opportunity to explore, exchange ideas, make new friends, and best of all come back next year to contribute future ComManTel conferences.

Honorary General Chair
Nguyen Minh Duc, REV Vice President, Chairman of ITC College of HCM City, Vietnam

General Chair
Khaled BenLetaief, Hong Kong University of Science and Technology, Hong Kong, China

Honorary Technical Program Chairs
Sherman Shen, University of Waterloo, Canada)
Tran Xuan Nam, Le Quy Don Technical University, Hanoi Vietnam

Technical Program Chairs
Trung Q. Duong, Blekinge Institute of Technology, Sweden
Nidal Nasser, Alfaisal University, KSA
Yan Zhang, Simula Research Inc., Norway

Signal, Image and Video Processing II

Chair: Le Tien Thuong, *Ho Chi Minh City University of Technology, Vietnam*

Cloud Computing and Services

Chair: Nguyen Huynh Tuong, *Ho Chi Minh City University of Technology, Vietnam*

How can human communicate with robot by hand gesture ?

Thi Thanh Hai Tran
MICA International Research Institute
HUST - CNRS/UMI-2954 - GRENOBLE INP
Hanoi University of Science and Technology
1, Dai Co Viet, Hanoi, Vietnam
Email: thanh-hai.tran@mica.edu.vn

***Abstract*—The use of hand gestures provides an attractive alternative to cumbersome interface devices for human - machine interaction (HMI). However, recognition of hand gestures is not a simple problem. In this paper, we propose to decompose the hand gesture recognition problem into 2 steps. In the first step, we detect skin regions using a very fast algorithm of color segmentation. In the second step, each skin region will be classified into one of hand posture class using cascaded Adaboost classifier and shape analysis techniques. The contribution of this paper is twofold. First, we proposed using both techniques for hand gesture recognition that reduces significantly the computational time in comparison with the traditional use of cascaded Adaboost classifier. Secondly, we integrated successfully this method on the robot and validated it in the context of interaction between human and robot guide in museum.**

I. Introduction

Robots are poised to fill a growing number of roles in today's society, from factory automation, service applications to medical care and entertainment. With the increasing of technology, scientists create new ways to make robots more and more intelligent that can replace human in dangerous jobs and situations and maybe make human everyday life a little easier.

While robots were initially used in repetitive tasks where all human direction is given a prior, they are becoming involved in increasingly more complex and less structured tasks and activities, including interaction with human required to complete those tasks. With this reason, Human - Robot Interaction (HRI) becomes a hot topic, that interests not only academic but also popular culture researchers. They want to study on how humans interacts with the robot, and how best to design and implement robot systems capable of accomplishing interacting tasks in human environment in a *direct, safe and effective* manner.

Ideally, human - robot interaction should approach human - human interaction which composes of many forms of communication. The main questions are: what kind of sensors can robot have can mimic human senses and how can robot replicate verbal and non-verbal communication using such sensors ? In early day of robots, researchers have been attempting to make them understand human speech. But in the last several year, researchers try to introduce the other means of human-to-human interaction to the field of HRI: hand gesture.

Human hand gestures are means of non-verbal interaction among people. Hand gestures have been shown to be an *intuitive and efficient* manner of communication. They range from simple actions of pointing at objects to the more complex ones that express our feelings or allowing us to communicate with others.

To make robot be able to communicate with human by hand gesture, it is essential that robot needs to be equipped with a camera to see human hand gesture and with "a brain" to analyze and understand which hand gesture in the captured image. Visual interpretation of hand gesture carries a tremendous advantage over other techniques that require the use of mechanical transducers: it is non-obstructive. There are few restriction imposed on the user's movement but visual interpretation also carries a burden of complexity in implementation due to variable illumination, cluttered background, large variability of hand's pose and scale, etc.

In this paper, we would like to demonstrate that: (1) in spite of the fact that Vietnamese people use rarely hand gestures in communication in general, hand gesture is a good manner for human - robot interaction and (2) the use of hand gesture to communicate with the robot is technically feasible. This is carried out by introducing a framework for designing a hand gesture vocabulary and a method for hand gesture recognition. We aim at providing (1) a *generic* set of hand gestures that can be used for many applications of service robot and (2) a *simple and efficient* method for hand gesture recognition, that could make the robot quickly understand the human intention in the communication in order to give suitable responses.

The paper is organized as follows. Section II depicts an overview of related works and introduces our contributions. Section III describes the general framework for human robot interaction and detail each component in the framework. Section IV shows experimental results of human robot interaction by hand gestures in real context. Section V concludes and gives some ideas for future works.

978-1-4673-2088-7/13/$31.00 ©2013 IEEE

II. Overview and Contributions

A. Related works

An increasing number of robotic systems are equipped for interaction with human users, each robot being designed for a specific objective. This specific design often limits theirs ability to perform complex human-robot interaction. Althrough there exists a large number of works on hand gesture recognition [1], [2], [3], [4], [5], the number of robots being able to perform advanced interaction due to multimodal perception with vision based gesture recognition is quite little. In this section, we present some examples of robots using hand gesture in communication with the users.

The Institute of Computer Design and Fault Tolerance developed a robot assistant, called Albert, that can be controlled by hand gesture and speech [6]. The robot is equipped with a color stereo camera allowing to recognize six types of hand gestures in order to stop the robot action in progress, to answer yes / no to the robot, to point to an object of interest or to explain grasp situation. The authors have validated their recognition algorithm but not experimented the communication between human and robot in a real situation.

The Faculty of Technology, Applied Computer Science, Bielefeld University developed the robot Biron as a companion of human in his household environment [7]. This robot can communicate with user through mono-manual pointing gestures. As designed to be a companion of human, for the first time, the user must give robot the tour through his private home so to familiarize it with its new habitat. The human points to and names locations and objects which she believes are necessary for the robot to remember. In the communication with human, only one type of hand gesture has been used.

The Institute for Computer Science of University of Bonn has a long history in developing generations of humanoid robot that can play soccer in the RoboCup Humanoid League and perform domestic tasks in RoboCup@Home League. In [8] two types of hand gesture showing and pointing in order to draw the robot's attention to a particular object have been studied. A Time-of-Light camera has been used for perceiving gestures. For pointing gestures, the pointing direction is estimated and matched with objects in the robot's environment. In the case of showing gesture, the robot tries to extract and visually recognize the shown object. One of disadvantage of this approach is the use of range camera that is not popular and quite expensive.

The group "Robotics and Interactions" (RIS) of LAAS has been conducting research on the autonomy of machines that integrate perception, reasoning, communication, learning, action and reaction capabilities [9], [10]. The Jido robot can interact with human user by both speech and gestures. Specifically, twelve types of hand gestures (mono, two handed gestures) including symbolic and deictic gestures have been used to communicate with the robot.

From these examples of robot, we found that:

- Each robot has been designed to perform a specific task (e.g. guide in the exhibition, companion in the house) so the type and the number of hand gestures used in communication with the user can be very different;
- Hand gesture type has been defined by robot designers, some gestures are common among applications (e.g. waving, pointing), some others have different meaning even the hand movement still remains the same.

In all cases, to make human robot interaction successful in reality, robot designers must care the problem of creating hand gestures set which are easy to perform by human and to recognize by the robot.

B. Contributions

In this paper, we present a framework for hand gesture based human robot interaction that can be integrated inside the interactive robot. It includes the study of Vietnamese's behavior when interacting with the robot then the definition of a vocabulary of hand gesture for communication and the proposition of a simple but efficient method for hand gesture recognition. This framework has been validated in a real environment with Vietnamese users, who are using less hand gesture in communication than others people in the world, that motivates the use of such modality of communication for any robot acting as an assistant, mostly in the context of Vietnam.

In comparison with our previous works [11], [12], this paper has two new points: (1) we propose to combine two techniques: shape analysis and cascaded adaboost classifier to improve performance of hand gesture algorithm in computational time (our previous work used only cascaded adaboost classifier for hand gesture recognition) and (2) we validate this framework in real context of human robot interaction (our previous work validated only the hand gesture recognition module by testing data captured a prior).

III. Human robot interaction by hand gesture

A. General framework for human robot interaction by hand gesture

To make the robot understand human hand gesture in the communication, a hand gesture vocabulary needs to be defined and learned by both human and robot and a hand gesture recognition needs to be developed and integrated within the robot. We divide these works into two phases: online and offline (Figure 1).

- Offline phase: This phase consists of (i) analyze human behavior in interacting with the robot; (ii) design hand gesture vocabulary; (iii) build hand gesture database and (iv) train hand gesture classifiers.
- Online phase: This phase will be activated at run time when human and robot communicate together. It is composed of (i) pre-process image; (ii, iii) detect and localize hand; (iv) recognize hand gestures.

In the following, we will describe in detail each components of each phase in the framework.

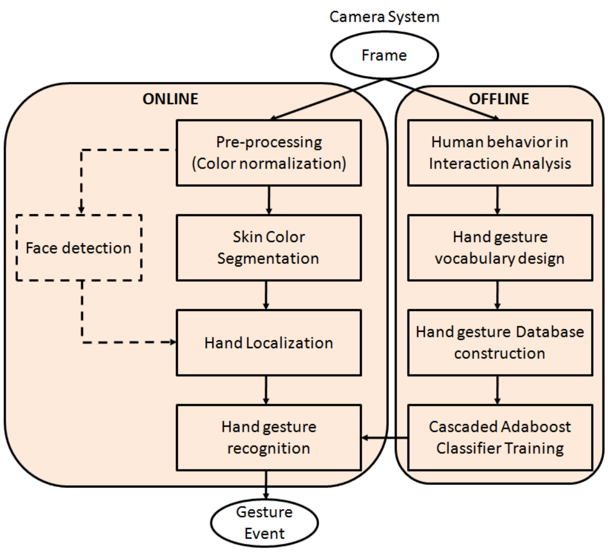

Fig. 1. Framework of hand gesture based human robot interaction proposed in this paper.

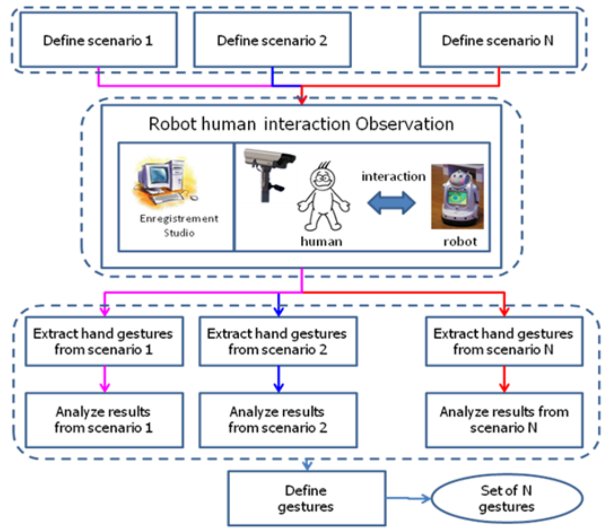

Fig. 2. Steps of hand gesture vocabulary design

Fig. 3. Set of five hand gestures used in situations of human - robot interaction.

B. Human behavior analysis and Design of hand gesture vocabulary

As mentioned above, to make human robot interaction successful in reality, robot designers must care the problem of creating hand gestures set which are easy to perform by human (*comfortableness*) and to recognize by the robot (*recognizability*).

A point that makes an important difference of our work in comparison with state of the art works is that we conducted the analysis of the use of hand gesture in communicating with the robot using Wizard of Oz technique - a famous technique for behavior analysis. By this way, we can find out hand gesture types which are often used by people in communicating with the robot, so the comfortableness can be assured.

The design of hand gesture vocabulary composes of 4 steps: (1) definition of interaction scenarios; (2) human robot interaction (HRI) observation in each scenario by camera; (3) hand gestures extraction and analysis; (4) definition of hand gestures set (Figure 2).

The result of this work is a set of five hand gestures which are distinct and used in almost interaction situations: (1) Call the robot to come for a service; (2) Point to an object of interest; (3, 4) Agree and Disagree with robot's answer; (5) Stop communicate with the robot (Figure 3) (see [12] for more detail).

C. Hand gesture database construction

In reality, there exists some hand gestures databases. However, these databases have some limitations: (1) the number of hand gestures and the number of people participating to build database are not large enough; (2) some databases consist only of gray scale images so can not be suitable for color based recognition techniques; (3) there are not exist a database of hand gestures of Vietnamese people. This is why we need to build a database of hand gestures to train and test our hand gesture recognition method.

We use two cameras to observe hand gestures of human in communication with the robot in frontal and profile views. Twenty people (ten men and ten women) are asked to perform five types of hand gestures under neon lighting condition, three times per gesture on two backgrounds (uniform and complex) (Figure 5). Finally we obtain 1200 videos, each has five second of duration. These data will be used for training and testing our algorithm. Later, they could be shared for research purpose in the domain.

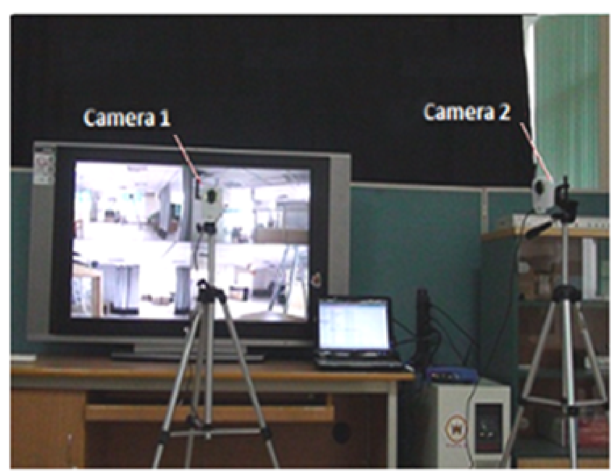

Fig. 4. Equipment preparation for collecting data.

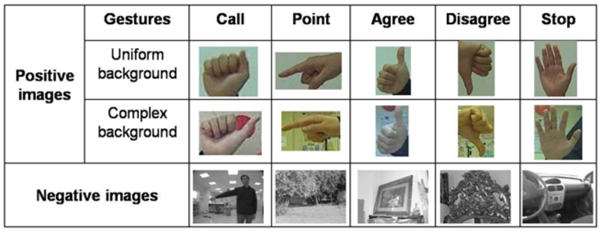

Fig. 5. Illustration of collected data for training and testing.

D. Method for hand gesture recognition

1) Hand detection using skin color segmentation: Hand detection is sometime not necessary but to discard rapidly non-interest regions, people often detect hand regions before recognize hand gesture. We propose to detect hand region based on skin segmentation. The skin segmentation can be done fastly with thresholding technique. In our work, by experience, we choose the most convenient thresholds that a skin pixel must satisfy: $(R > G)AND(R > B)AND(R > 95)AND(G > 40)AND(B > 20)$.

The Figure 6 shows an example of image segmented by skin thresholding technique. We can see that skinlike pixel have been displayed as "white" pixels in the image. By this way, face and other skinlike regions (office table) are also detected. We need to discard faces using face recognition [13] and skip small regions before applying hand gesture recognition.

Fig. 6. Skin based hand segmentation. In this figure, face and hand are segmented and displayed as white pixels.

2) Hand posture classification: To classify a hand region into one of six categories $\{$*Call, Point, Agree, Disagree, Stop, other*$\}$, we propose an approach that combines the result of recognition from shape analysis and cascaded adaboost classifier. In the literature, shape analysis and cascaded adaboost classifier are widely used for hand gesture recognition but there are no work on combining two methods. Cascaded adaboost classifier must scan all candidate windows in the image and verify if it contains a hand gesture. However, the number of windows is very large and almost of them are background. Discarding non-hand regions will speed up hand gesture recognition.

- **Shape analysis:** Once hand is detected in the image, skin regions containing hand will be determined by binarization algorithm. The shape features such as elongation, compactness will be calculated from the region and used to classify hand posture.
- **Cascaded adaboost classifier:** Shape based classifier is simple and fast. In addition, it is invariant to translation, rotation, scale change. However, this method gives quite frequently false alarms. Therefore we propose to combine it with a stronger classifier based on boosting approach. Each posture is trained with a cascaded adaboost classifier using Haarlike features (see our previous work [11]). The structure of cascaded adaboost classifier is illustrated in Figure 7.

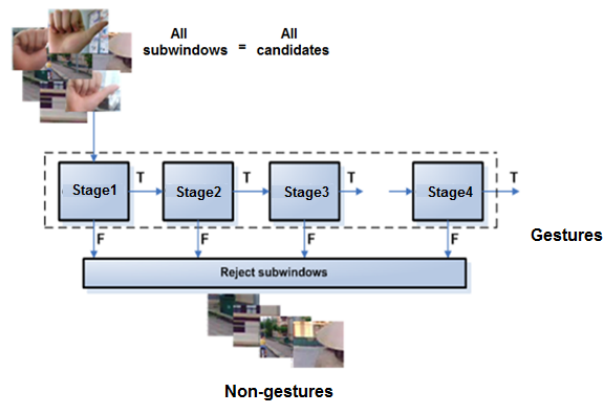

Fig. 7. Structure of Cascaded Adaboost Classifier.

- **Result fusion:** We simply fuse the results of both approaches as following: For each skin windows, it will pass all five cascaded adaboost classifiers and shape based classifier. Each cascaded adaboost classifier will give a response 1 or 0 corresponding to gesture or non-gesture while shape based classifier gives one type of hand gesture. The candidate window will be classified into one of six categories if it is well classified in this category by both shape based classifier and one among five cascaded adaboost classifiers (Figure 8).

IV. Experiments

A. Context of experiment

1) General description: We will validate our hand gesture recognition algorithm in the context of human robot interaction. The hand gesture recognition module will be integrated on the robot 914 PCbot. Camera is installed at one meter of height, so that can observe human standing at long distance (1.5-2m).

2) Definition of experimental scenario: We test the human robot interaction by hand gesture in two environments: Showroom of our institute and exhibition room of Ethnology Museum of Vietnam. We propose a scenario in which the robot

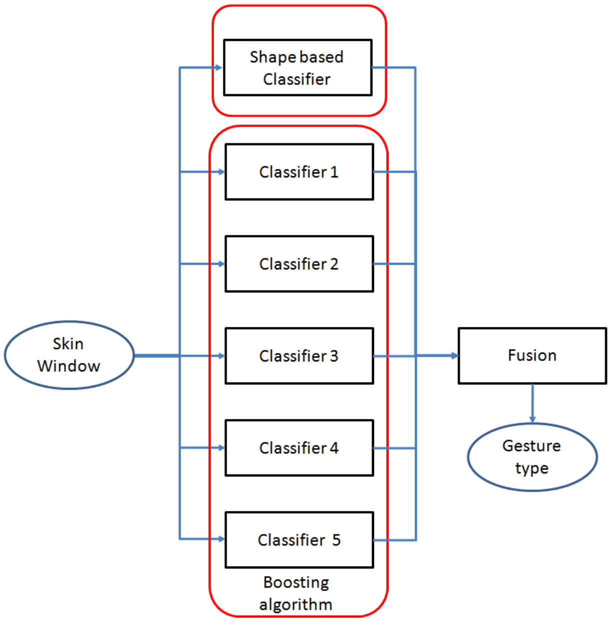

Fig. 8. Fusion of recognition results.

Fig. 9. Example of hand gesture recognition.

is active to communicate with users. Once it detects faces in its camera's field, the robot will communicate with the person having the biggest face. The robot says *Hello* to the human by synthesis speech then proposes some services to him like presentation about the robot, presentation of objects in the museum. If the robot can not detect face, human can use *Call* gesture to say hello and call the robot to come for asking a service. The human can point to an object in the museum and ask information about this object. Human expresses his attitude to the robot through hand gestures *agree, disagree*. Robot gives suitable answers (*Sorry* in case of disagree and *Thank you* in case of agree) by synthesis speech. When all information are provided, the user can stop the communication using *Stop* hand gesture.

B. Experimental results

We experiment the same scenario in two cases. The first case takes place at Showroom of our institute (neon lighting condition, office background) with ten people (five men and five women), each perform ten times the scenario in which four people have participated to build training database. The communication is fluent and the recognition rate of hand gestures is about 77 % (the same recognition rate obtained with cascaded adaboost classifier) at 45 milliseconds per frame on the PCbot (2G RAM, Processor T7200 2GHz). We can observe a significant improvement of speed with respect to the method based only on cascaded adaboost classifier (150 milliseconds per frame).

The second case takes place at Museum. Visitors can communicate easily with the robot by hand gesture and speech. In museum, the lighting condition is difficult due to the variation of light source, however, our proposed hand gesture recognition still worked. The figure 10 illustrates the robot in communication with a visitor in the museum.

Fig. 10. Human robot interaction in museum.

V. Conclusions

In this paper, we showed that hand gesture is an effective manner to communicate with the robot. A new point in this paper is we combined skin based hand detection and shape analysis with cascaded adaboost classifier. By doing it, the speed of recognition becomes very fast even on a less powerful machine. This improvement brings a significant advantage mostly when we would like to build interactive robot which communicates with human not only by hand gestures but others modalities, this requires resource for each modality. In the future, we want to evaluate our framework in a real situation (communication between human and robot in a museum with a big enough number of participants). In addition, we will consider more types of hand gestures, even dynamic hand gestures so that the communication between human and robot becomes more and more natural.

Acknowledgment

This work is partially supported by the research project *"Visually impaired people assistance using multimodal technologies"*, funded by the VLIRs Own Initiatives Programme, under the grant reference VLIR-UOS ZEIN2012RIP19 and by the project *"Study and develop an advertising system in online image sharing services"* funded by Vietnamese Ministry of Science and Technology - 2012.

References

[1] W. T. Freeman and M. Roth, *Orientation Histograms for Hand Gesture Recognition*, in Proc. of International Workshop on Automatic Face and Gesture Recognition, pp. 296-301, 1994.

[2] K. Nickel and R. Stiefelhagen, *Visual recognition of pointing gestures for humanrobot interaction*, Image and Vision Computing, 2006.

[3] C. Wang and K. Wang, *Hand Posture Recognition Using Adaboost with SIFT for Human Robot Interaction*, International Conference on Advanced Robotics, 2007.

[4] Y. Fang, J. Wang, K. Wang, J. Liu and H. Lu, *Hand Posture Recognition with Co-Training*, in 19th International Conference on Pattern Recognition. 2008. Tampa, FL.

[5] H. Tran, *Vision based hand gesture recognition: a survey*, Vietnamese Journal on Information and Communication Technology, 2011

[6] O. Rogalla, M. Ehrenmann, R. Zollner, R. Becher and R. Dillmann, *Using speech and hand gesture control for commanding a robot assistant*, In Proc. of 11th IEEE International Workshop on Robot and Human Interactive Communication (2002)

[7] J. F. Maas, T. Spexard J. Spritsch, B. Wrede and G. Sagerer, *BIRON, What's topic ? A multimodal Topic Tracker for improved Human Robot Interaction*, In Proc. of The 15th IEEE International Symposium on Robot and Human Interactive Communication, UK, Sept. 2006.

[8] D. Droeschel, R. Stuckler, D. Holz and S. Behnke, *Towards joint attention for a domestic service robot - person awareness and gesture recognition using time-of-flight camera*, in Proceeding of IEEE International Conference on Robotics and Automation, Shangai, China, 2011.

[9] A. K. Pandy and R. Alami, *Towards task understanding through Multi-state visuo-spatial perspective taking for Human-Robot Interaction*, In Proc. of International Joint Conference on Artificial Intelligence (IJCAI'11), Barcelone (Espagne), 16-17 Juillet 2011, 6p.

[10] B. Burger, I. Ferranee, F. Lerasle and G. Infantes, *Two-handed gesture recognition and fusion with speech to command a robot*, Autonomous Robots, 2012, Vol.32, N2, pp.129-147.

[11] M. Nguyen, H. Pham, T. Dong, S. Nguyen and H. Tran, *A fully automated hand gesture recognition system for human robot interaction*, In Proc. of the Second International Symposium on information and communication technology SoICT 2011, ISBN: 978-1-4503-0880-9, pp.112-119

[12] M. Nguyen, H. Tran and S. Nguyen, *Wizard of Oz technique for Designing Hand Gesture Vocabulary in Human Robot Interaction*, the third International Conference on Knowledge and system Engineering, KSE2011, Publised by IEEE Computer Society, ISBN 978-0-7695-4567-7, pp.232-237

[13] P. Viola and M. J. Jones, *Robust Real-time Object Detection*, Cambridge Research Laboratory Technical Report Series, 01/2001.