行政院國家科學委員會專題研究計畫 期中進度報告

家庭全方位之多樣性群組型娛樂/休閒機器人--子計畫四：家庭成員與娛樂休閒機器人之行為互動智慧型控制(1/3)
期中進度報告(精簡版)

計 畫 類 別 ：整合型
計 畫 編 號 ：NSC 96-2218-E-002-030-
執 行 期 間 ：96 年 08 月 01 日至 97 年 07 月 31 日
執 行 單 位 ：國立臺灣大學電機工程學系暨研究所

計 畫 主 持 人：連豊力

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 97 年 05 月 30 日

行政院國家科學委員會專題研究計畫期中進度報告

家庭全方位之多樣性群組型娛樂/休閒機器人

子計畫四：家庭成員與娛樂休閒機器人之行為互動智慧型控制(1/3)

計畫編號： NSC 96-2218-E-002-030-
執行時程： 96 年 8 月 1 日至 97 年 7 月 31 日
計畫主持人： 連豊力　國立台灣大學電機系 副教授
參與人員： 林意淳、周執中、曾冠傑、林宇騫

中文摘要

本計畫第一年度的研究課題主要著重於具備社會互動能力的家庭休閒娛樂型機器人設計。具備社會互動能力的基本精神乃是要在日常的家庭生活中，機器人能夠是適切地輔助與強化家人間的活動，這類機器人的設計主要必須包含多模式的認知與感官功能，以便於瞭解家人的情緒反應，行為模式，以進而採取適當的互動反應，在本年度，我們規劃了一個家長與機器人在小孩成長學習過程中的一個互動機制模式，同時，測試了家人之間以及人類與機器人之間的互動法則。

關鍵字: 社會互動、家庭服務機器人、人機互動

***Abstract*—The applications of robot system include assistance robots, service robots and health care robots. In this report, the home service robots with social interaction capability are mainly focused. Home service robots with social capability are expected to assist people in their home-based activities. In order to implement the social interaction with human, the robots should involve the following three key modules: multimodal, cognition and emotion module. The multimodal module consists of input, output and fusion components. A general human-robot interaction framework is described in detail. An example of the simple interaction structure of robot-child-parent and the experiment scenario are presented.**

***Index Terms*—social interaction, home service robot, human robot interaction**

I. INTRODUCTION

THE development of robots has been tended to serve human in industry and factory environment. Therefore, it is an important issue on how to design robots to perform home service and provide all kinds of entertainment activities to human. When a robot becomes a part of human daily life, the interaction relation between human and robot is important. The subject has been a key research topics in robotics and other research fields, including psychology, artificial intelligence, and human-robot engineering.

In recent years, the role of robots has been changed from an instrument to a partner or a friend of human in the workplace or family. Meanwhile, researchers attend to study the way to change the emotion of a robot, and to establish the complex social relationship and cooperation of human with robots. The research topics are used to understand the two-way social interaction between human and robots. The fundamental of these researches is mainly based on psychology and biology.

Strictly speaking, the interaction relationship between human and robot is not identical with artificial intelligence. The artificial intelligence is to minic the expression of human directly. However, the human-robot interaction is to analyze and comprehend the social interaction in intelligence and to build a suitable respond pattern. Moreover, the human-robot interaction not only studies the human intelligence, but also discusses interactively learning model.

Traditionally, the design of robot only considers the interaction between robot and environment in one pre-planned task. Hence, the robot by a traditional design can not point out clearly the interaction content and timeing between human and robot. Therefore, it is required to establish a common ground between human and robot by analyzing their interaction pattern. The basic design principle is to integrate all sensory data then to combine the cognitive process and emotional expression. Besides, it also employs the self societal to find the interaction framework between human and robot [1: Huttenrauch et al. 2004].

The human usually uses the body language to support the voice and only needs the classification of quality in cognitive capability. However, the cognitive capability of robot needs more specificity quantification standard. For example, the representation of the red color is distinct in difference human. The representation is a rough description. But, in a robot system, the values of RGB are used to be a quantified data to represent the red color. In consequence, the human-robot interaction model should have the capability of management two different cognitive types.

The study on the emotion of robot is based on the psychology.

Contact Information: Feng-Li Lian, phone: 886-2-3366-3606; e-mail: fengli@ntu.edu.tw.

The robot is proposed by Breazeal et al. can determine the direction of eyes, the facial expression and the voice recognition. It also can show several kinds of emotion state such as happiness, sadness, smile. Furthermore, this robot is easy to make a communication with human [2: Breazeal 2003].

Several applications related to the development of interaction between human and robot are summarized below. Ahn et al. make a photo-robot which can interact with human as well as a cameraman. The main function of photo-robot is to determine the action of wave and approach human to take a photo and upload the photo into the personal computer through wireless network [3: Ahn et al. 2006]. Austermann et al. develop a robot named MEXI which can control facial components to respond by determining the emotion of human [4: Austermann et al. 2005]. Song et al. research a robot system can track and orientate the position of eyes in a noise environment by determining the direction of sound [5: Song et al. 2006]. Kato et al. present a cute robot named Ifbot which has various facial expressions to communicate with human [6: Kato et al. 2004].

The rest of the report is organized as follows. Section II discusses the related work of interaction between human and robot. Section III presents the human-robot interaction framework which includes multimodal nodule, cognitive module and emotional module. Section IV illustrates the experimental design in family. Section V summarizes this report.

II. Related Work

In the last few years, several articles devoted to the study of social interaction between robot and human. Besides, the social interaction between human and human has been researched by sociologists. Robot communicates with human by understanding the gestures and the emotions of human. Therefore, it is important to understand the social interaction behaviors of human with human and human with robot. Two main survey directions for the interaction behavior are engineering and sociology. The purpose of sociological survey is to design a suitable interaction architecture in the engineering based on the sociological theories for human and robots. Related literature survey is discussed in detail in the following.

A. Engineering

First, use keywords such as social interaction and human robot interaction, separately, and search for related papers within last two years in the website of IEEE Xplore. According to different conferences and years, the amounts of papers are summarized in Table 1. Most of the papers related to the interaction between human and robot were appeared in the IEEE/RSJ International Conference on Intelligence Robots and Systems (IROS) and IEEE International Workshop on Robot and Human Communication (ROMAN).

The intelligence robot system includes four types: (1) service robots (2) assistance robots (3) social robots, and (4) health care robots. In this report, social robots are mainly focused in detail. The study of social interactions can be explained in terms of the relation between human and robots, the using duration of robots by human, and the purpose of using robots. The relationship between people and robots can be divided into active and passive. The using duration of robots by human can be categorized into long-term and short-term. For short-term, the purpose of using robots is to provide services, information and entertainment. For long-term, the purpose is to exhibit itself personality and character. Robot owns the capability of facial expression, speech recognition, movement, emotional, and social contact [7: Gockley et al. 2005]. As mentioned above, these are summarized in Table 2.

Table 1. The number of paper about social interaction in different conference.

Conference (year)	Numbers
IROS(2007,2006)	48
RoMan(2006,2005)	32
ICRA(2007,2006)	23
IECON(2007,2006)	9
SICE(2007,2006)	5

Table 2. The type of social robot

People	Robot Type	Purpose	Equipment
Do not Actively interact with robot	Short-term Nonverbal Interaction	Not provide information	Facial Expression Movement
	Short-term	Serve, Provide Information And entertain	Speech Facial Expression Emotional
Actively interact with robot	Long- term	Service Exhibit Personality And character	Social competence

The position of human is divided into bystander and participant in a communication. When a robot becomes as medium between human and human, the type of medium can be categorized into passive, interactive, and passive-social as shown in Figure 1. The passive medium means that the robot merely provides information to people. The interactive medium means that the robot not only receives requests from people but also presents information to people. The passive-social medium means that the robot does not accept requests from people, but presents more information to people through its social ability which is the expression of conversation. The application case of passive-social is shown in Figure 2. Two robots perform the Manzai show which is Japanese comedy conversation by signal exchange.

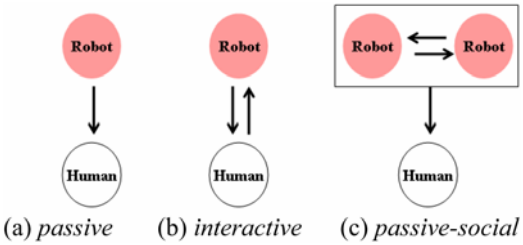


Figure 1. Three types of medium of robots. [8: Hayashi et al. 2005]


Figure 2. Two robots perform a Manzai show in Japan. [8: Hayashi et al. 2005]

(a) (b)

Figure 3. (a) A receptionist robot of facial expression, speech, and background storytelling. [7: Gockley et al. 2005]
(b) A robot interacting with onlooker. [9: Hollinger et al. 2006]

The application case of long-term in Figure 3(a) is to install a receptionist robot in the front door at the Robotics Institute and the School of Drama Carnegie Mellon. The receptionist robot has the capability of facial expression, speech, and storytelling. Moreover, the receptionist robot also receives the request from user by keyboard and provides information to user. Figure 3(b) shows the differential-drive RWI Magellan Prorobot to interact with onlooker by detecting the color of shirt of the onlooker. The Magellan Prorobot plays different music to people according to the color of shirt of onlooker.

B. Sociology

The main survey of sociology is from the book. Use keywords such as social interaction and role theory, separately, and search for several related books. One of the early social interaction process models is composite of Mead's and Schutz's model. The composite of Mead's and Schutz's is composed of self, generalized framework, and signal. The self includes three parts which are self reference, deliberative capacities, and stocks of knowledge. The generalized framework is a social code of conduct. The signal contains the signals of self and others, and how to interpret. The detail interaction processes between self-reference, deliberative capacities, and signals are shown in Figure 4. In other words, the human interaction is possible because people assume that they have common stocks of knowledge and use these stocks to adapt themselves in time and space. Then, humans use knowledge to explain the contextual meaning of gesture, categorize objects and people, and determine the appropriate rules or procedures for making inferences about the signal of others.

The entertainment in the family is focused mainly in the paper. For instance, the interaction process between parent and child can be regarded as basic process. Therefore, using keyword as family or art education of child search related book.

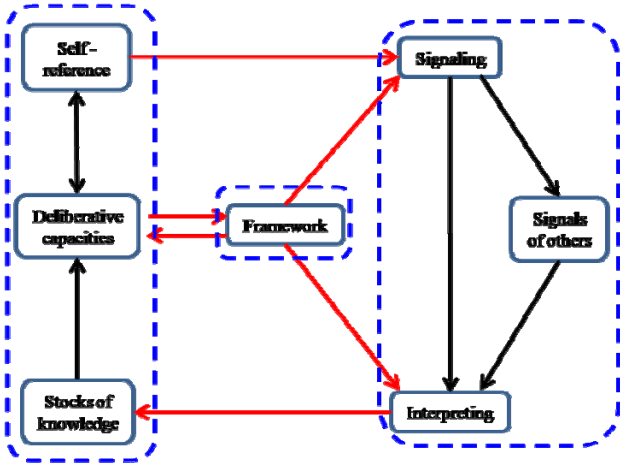
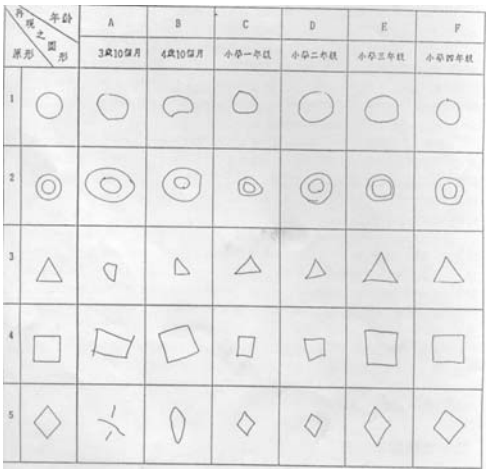


Figure 4. Composite of Mead's and Schutz's model. [12: Turner 1998]

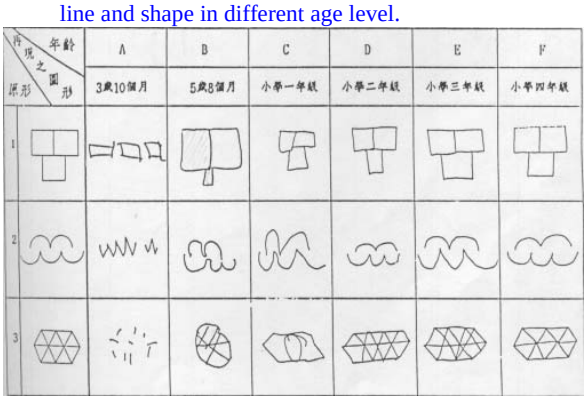
Piaget and Hebb propose the development of perception of child is stage. The experiments are designed based on above theory. The experimenters choose the children which age within three and ten years old. They have no the training about art before they go to elementary school. Besides, the grade of the drawing lies to the average in a class. The experimenters observe these children to understand the perception of eye through drawing. Table 3 shows the result of the children drawing the simple geometric as circle, triangle, square, and rhombus. Table 4 shows the result of drawing more complex shape. The two tables can used to establish a database of behavior index.

Table 3. Drawing the difference simple geometric in different age level.



[15: 台灣省國民學校教師研習會 1996]

Table 4. Drawing the difference complex geometric as the overlap of line and shape in different age level.



[15: 台灣省國民學校教師研習會 1996]

III. MULTIMODAL INTERACTION

In this section, the human-robot interaction framework is studied. The human-robot interaction framework is composed of three modules as multimodal interaction module, cognitive module, and emotional module. Three modules are related to each other and the division between modules is based on the level of information processing stages. The multimodal interaction module plays a role of interface to receive the varying inputs and expresses suitable actions. The cognitive module is more task-oriented and cooperates with user to achieve the goal of the task. The emotional module is more human-oriented and maintains the social relation. Figure 5 shows the human-robot interaction framework in the robot.

The multimodal interaction module is based on three parts as input, output, and fusion. The multimodal module not only simply means the number of sensory, but also refers the quality of the system. The multimodal module has the mediate interface as well as GUI. The mediate interface is a portable and has vocal interaction channel, touchscreen, speaker, and camera. The multimodal module also has a selective attention module. The purpose of selective attention module is to reduce the amount of information from inputs. The inputs can be weighted according to degree of importance and integrated into a unified input set after fusion [10: Lee et al. 2005].

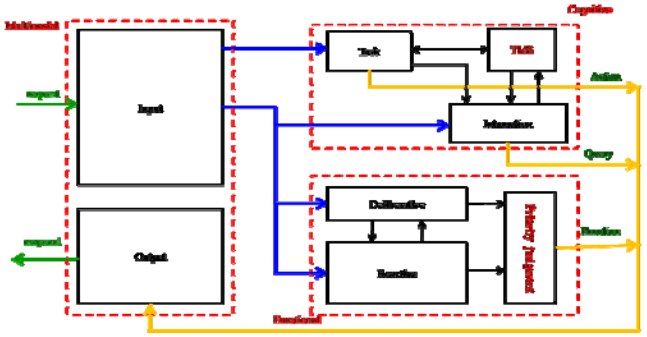


Figure 5. The human-robot interaction framework with multimodal module, cognitive module, and emotional module.

The cognitive module is consisted of task model, interaction model, need model and truth maintenance system (TMS). The purpose of task model is to generate the details of action procedures and modify the details according to the continual communication with the human. The interaction model handles the questions and suggestions from user. The need model includes the situations and requirements as well as commonsense restrictions and safety-related constraints. The truth maintenance system is used to continually update the correctness for the combination verbs and objects. Meanwhile, it reduces effectively the ambiguities of words.

Emotions are important in social interaction, whether verbal or non-verbal. Moreover, the emotions have extremely effect on cognitive process as decision marking, planning and learning. The emotional module is composed of deliberative model, reactive model, and priority judgment. The deliberative model supports indirect emotions. It provides emotional states to cognitive model and receives the goals of current task from cognitive model. The reactive model supports immediate emotional reaction. Using this mechanism to realize an entertainment robot more resemble a partner of human. The aim of priority judgment is to deal with conflicts between reactive layer and deliberative layer in an emergency situation.

For instance, in a family life, the members consist of father, mother, grandmother and two children, one of that is junior high school student and the other is elementary student. They all have a robot belong to themselves. Their robot can detect the environment and provide or record need of a daily life for everyone. Robots can communicate with each other and assist or entertainment human. In the next section, more details of experiment design are presented.

IV. EXPERIMENT DESIGN

According to the sociology theory has been mentioned in Section II, we apply it to the interaction in a family scenario. In Figure 6, parent, child and robot correspond with framework, signal and self of Composite of Mead's and Schutz's model. Because the interaction is implemented in robot, child means signal input and parent mean constraints for robot. The

existence of robot is a bridge for human. The purpose of the framework is to raise the convenience and satisfy various entertainment need to members of family. Figure 6 shows a simple interaction structure with robot, parent and child.

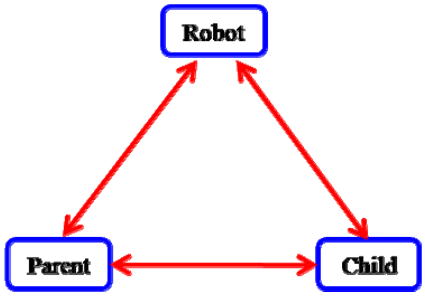


Figure 6. The interaction relation exists among robot, parent, and child.

The robot carries various scenarios such as playing with child, studying with child, watching safe of child and taking care of child and so on. The robot has multimodal interaction model, cognitive model and emotional model in every scenario. Figure 7 shows more details interaction process. Step 1: parent decide one scenario through touchscreen on robot. Step 2: robot adopts a suitable action and interaction to child and enhances the right behavior through detecting the respond or emotion of child. Step 3: based on the cognition and emotion level of child to analyze the responds and categorize those then robot responds to child. Step 4: when child finishes action or shows bad emotion, the robot informs parent of the time that they should appear.

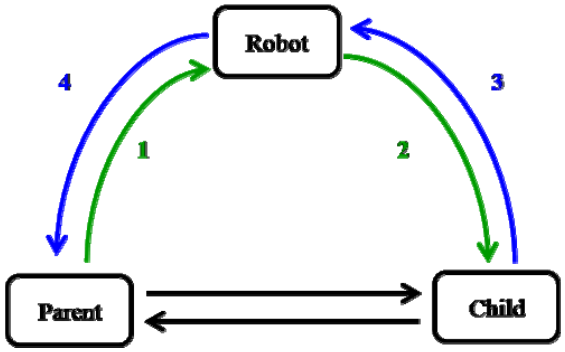
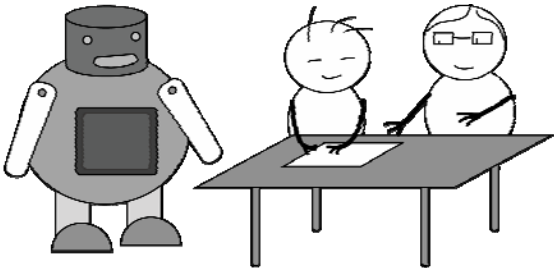
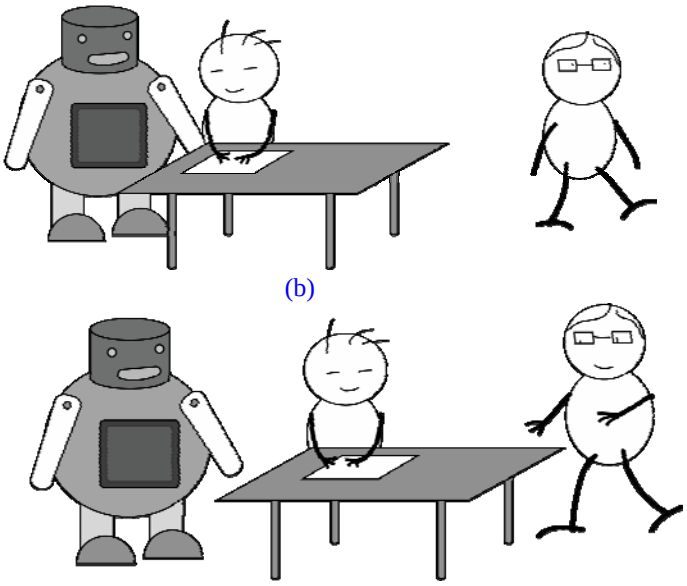


Figure 7. The interaction framework of robot, parent, and child.



(a)



(b)

(c)

Figure 8. Learning scenario.
(a) Robot observes the interaction of father with child.
(b) Robot becomes as medium between father and child.
(c) Robot informs the father appearance.

Figure 8 shows a learning scenario. First, father is teaching child how to draw a triangle and robot only observes the interaction between father and child. When child starts to ask repeated actions or questions and father has to do other thing as job, the robot becomes a bridge or medium between father and child. The robot has the capability of teaching child to draw a triangle. The communication channel of robot and father uses network. When emergency situation happens, robot sends a message to inform father. It is important that the robot is not to replace the position of father. The design of robot is to create a better entertainment in family.

V. Conclusion

The development of robot toward the social interaction has been studied in last decade years. In engineering, the research is focus on the hardware equipments as moving, sensory and expression. The social interaction between robot and human is pre-planned algorithm. In sociology, interaction process models have been proposed and one of those is used to establish our interaction model in family. In order to adapt to the various environment and responds from people, the multimodal interaction module is necessary. Further, cognition and emotion process are important component of interaction behavior between human and human. Hence, robot also has cognitive and emotional module. Section VI presents a experiment scenario about the interaction among robot, child and parent.

The future work is to analyze the appearance time and stay time of father in interaction process. To characterize the quantity and quality of symbol transmitted by network and to correspond the meaning of symbol to words used in human interaction.

VI. 計畫成果自評

- 查核點及審查要項：

查核點		查核項目	具體評量指標
第一年度	上半年	休閒娛樂的需求與影響	完成居家環境之中，不同年齡層的成員在於休閒娛樂的需求型態與影響規劃
			分析在家庭中，兒童、青少年、中年人、老年人，對休閒娛樂的需求型態以及所具備的影響性。
		人類互動行為模式與訊息交換機制	完成對家庭中的成員於互動上所需要之基本行為模式與相關訊息交換機制之設計與規劃
			在社會學中，存在著討論人類社會互動結構的模型，我們利用這個模型設計出在家庭中，父母、機器人、小孩的基本的互動行為模式。關於訊息交換的機制主要是建立在，手勢、語言、音調、情緒和個人認知背景，這五個重要的因素上。進一步定義出在每次的互動中，訊息所代表的量與質，接著歸納出數個層級，並且實際應用所設計的情境例子中。

查核點		查核項目	具體評量指標
第一年度	下半年	人機互動溝通平台	完成一套家庭成員與各類機器人之間的互動溝通平台之設計
			對於家庭中人機互動溝通平台，藉由人與人互動的模式架構，推展至人機互動溝通的平台。對於平台所採用的網路協定，討論在傳送符號訊息的過程中，符號的量與質與實際人類互動所採用的訊息的相對應。當機器人變成人與人溝通之間的媒介時，傳遞訊息的頻率就顯得很重要，藉由Shannon's information theory，來分析傳遞訊息所需頻寬與實際頻寬之間的關係。
		行為模式基本法則	完成行為模式機制之基本法則包括： 多模式行為、認知互動、以及情緒化互動
			對於多模式行為主要分為輸出和輸入兩部份，輸入是指人的情緒、姿態、聲音、詞語等的表達，經由資訊融合分析後，再送到認知互動和情緒互動的模組，輸出則是指機器人的反應，根據認知與情緒互動模組的處理結果，來決定採取那種反應。認知與情緒的互動模組，需要先建立各種行為所對應的情緒反應的資料庫，作為比對的參考，在認知方面，要有工作任務的排程的能力，來協助人類達到所想要的目標，在情緒方面，要有優先權調整的機制，可以處理緊急情況的發生。

REFERENCES

[1: Huttenrauch et al. 2004]
H. Huttenrauch, A. Green, M. Norman, L. Oestreicher, K.S. Eklundh, "Involving Users in the Design of a Mobile Office Robot," *IEEE Transactions on Systems, Man and Cybernetics, Part C*, vol. 34, no. 2, pp. 113-124, May 2004.

[2: Breazeal 2003]
C. Breazeal, "Emotion and sociable humanoid robots," *International Journal of Human-Computer Studies*, vol. 59, no. 1-2, pp. 119-155, July 2003.

[3 :Ahn et al. 2006]
Hyunsang Ahn, Dohyung Kim, Jaeyeon Lee ,Suyoung Chi, Kyekyung Kim, Jinsul Kim, Minsoo Hahn, and Hyunseok Kim, "A Robot Photographer with User Interactivity," *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5637-5643, Beijing, China, Oct. 2006.

[4: Austermann et al. 2005]
Anja Austermann, Natascha Esau, Lisa Kleinjohann and Bernd Kleinjohann, "Prosody Based Emotion Recognition for MEXI," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1138-1144, Aug. 2005.

[5: Song et al. 2006]
Kai-Tai Song, Jwu-Sheng Hu, Chi-Yi Tsai, Chung-Min Chou, Chieh-Cheng Cheng, Wei-Han Liu, and Chia-Hsing Yang, "Speaker Attention System for Mobile Robots Using Microphone Array and Face Tracking," *Proceeding of the 2006 IEEE International Conference on Robotics & Automation*, pp. 3624-3629, Orlando, Florida, May 2006.

[6: Kato et al. 2004]
Shohei Kato, Shingo Ohshiro, Hidenori Itoh, Kenji Kimurai, "Development of a Communication Robot Ifbot," *Proceedings of the 2004 IEEE International Conference on Robotics & Automation*, pp. 697-702, New Orleans, LA, April 2004.

[7: Gockley et al. 2005]
R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A. C. Schultz, and J. Wang, " Designing Robots for Long-Term Social Interaction," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005).* pp. 1338-1343, 2-6 Aug. 2005.

[8: Hayashi et al. 2005]
K. Hayashi, T. Kanda, T. Miyashita, and H. Ishiguro, N. Hagita, "Robot Manzai–Robots'Conversation as a Passive Social Medium," *Proceedings of 5th IEEE-RAS International Conference on Humanoid Robots*, pp. 456-462, 5-7 Dec. 2005.

[9: Hollinger et al. 2006]
G. A. Hollinger, Y. Georgiev, A. Manfredi, B. A. Maxwell, Z. A. Pezzementi, and B. Mitchell, "Design of a Social Mobile Robot Using Emotion-Based Decision Mechanisms," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China*, pp.3093-3098, 9-15 Oct. 2006.

[10: Lee et al. 2005]
K. W. Lee, H.-R. Kim, W. C. Yoon, Y.-S. Yoon, D.-S. Kwon, "Design A Human-Robot Interaction Framework For Home Service Robot," *IEEE Interactional Workshop on Robot and Interactive Communication*, pp.286-292, 2005.

[11: Murcie 2002]
John Murcie, *家庭社會學*, 韋伯文化國際, 2002.

[12: Turner 1998]
Jonathan H. Turner, *A theory of social interaction*, Stanford University Press, 1988.

[13: Bales 2002]
Robert Freed Bales, *Social interaction systems : theory and measurement*, New Brunswick, N.J. : Transaction Publishers, 2002.

[14: Hare et al. 2005]
A. Paul Hare et al., *Analysis of social interaction systems: SYMLOG research and applications*, University Press of America, 2005.

[15: 台灣省國民學校教師研習會 1996]
台灣省國民學校教師研習會, *兒童美術教育理論與實務探討*, 台灣省國民學校教師研習會, 台北縣板橋市, 1996.