Conversational Informatics and Human-Centered Web Intelligence

Toyoaki Nishida, *Member, IEEE*

***Abstract*— Conversation is the most natural communication means for people to communicate with each other. I believe that conversation plays a critical role in realizing a paradigm of human-centered web intelligence in which web intelligence engines are grounded on the human society. We are currently building a computational framework for circulating information in a conversational fashion, using information packages called conversation quanta that encapsulate conversational scenes. Technologies are being developed for acquiring conversation quanta on the spot, accumulating them in a visually recognizable form, and reusing them in a situated fashion. Conversational Informatics, based on measurement, analysis, and modeling of conversation, constitutes the theoretical foundation for these applications. I will overview recent results in Conversational Informatics that will help achieve our vision. I will also discuss our approach in the context of Social Intelligence Design aimed at the understanding and augmentation of social intelligence for collective problem solving and learning.**

***Index Terms*— Conversational Informatics, Social Intelligence Design, Human-centered computing, Human computer interaction**

I. Introduction

THE goal of Web Intelligence is to create a world wide wisdom web (w4) by integrating individual intelligences available on the global network using technologies such as web agents, web mining and farming, web information retrieval, web knowledge management, web intelligence infrastructure, and social network intelligence [1].

In order for Web Intelligence to be able to maximally benefit the human society, it should be well-interfaced to the human society so that each member of the human society can benefit from it without much difficulty and Web Intelligence can gain enough sources of knowledge from the human society. Web Intelligence will synergistically co-evolve with the human society if it is intimately embedded in the human society.

One of the key issues in embedding Web Intelligence in the human society is information grounding, which roughly means that the information user is aware of the association between information and the real world. Web Intelligence need to provide information in such a way that people can readily ground it on their daily life. Even though potentially useful information is provided with Web Intelligence, it might be useless if the information user fails to recover the reference to real world or reconfigure the image implied by the given statement. Unfortunately, information grounding is not easy to establish once information is isolated from the original situation it is created unless special care is taken at the moment information is created. We need to invent a technology that permits people to preserve cues for information grounding on the spot so that they can help ground the information later at the situations different from the original one.

In this article, I focus on the role of conversation in information grounding and present a suite of technologies aimed at realizing a paradigm of human-centered web intelligence. Apparently, conversation is the most natural communication means for people to communicate with each other. A closer look reveals that various kinds of processes related to creation or recovering information grounding are in action in conversation. For example, pointing and gaze are basic forms of creating association with the real world and co-occurring propositions. Gestures and postures may suggest the scope and modality of the utterances. In addition, dialectic aspects of conversation help participants interpret the meaning of information through discussions.

We are currently building a computational framework for circulating information in a conversational fashion, using information packages called conversation quanta that encapsulate conversational scenes consisting of participants' behavior, references to the environment, and meta-descriptions. Technologies are being developed for acquiring conversation quanta on the spot, accumulating them in a visually recognizable form, and reusing them in a situated fashion.

Conversational Informatics is a field of research aimed at establishing the theoretical foundation for these applications, based on measurement, analysis, and modeling of conversation. The field exploits a foundation provided by Artificial Intelligence, Natural Language Processing, Speech and Image Processing, Cognitive Science, and Conversation Analysis. It is aimed at shedding light on meaning creation and interpretation resulting from the sophisticated mechanisms in verbal/nonverbal interactions during conversation, in search of better methods of computer-mediated communication, human/computer interaction, and support for knowledge

Toyoaki Nishida is with Dept. of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan (phone: +81-75-753-5371; fax: +81-75-753-4961; e-mail: nishida@i.kyoto-u.ac.jp).

creation. I will overview recent results in Conversational Informatics that will help achieve our vision.

I will also discuss our approach in the context of Social Intelligence Design aimed at the understanding and augmentation of social intelligence for collective problem solving and learning.

II. Conversational Informatics

Conversational Informatics [2] is a field of research that serves as a theoretical ground for understanding conversational phenomena and developing conversational systems. By integrating the methods in Artificial Intelligence, Pattern Recognition, and Cognitive Science, Conversational Informatics addresses understanding and augmenting conversation. Currently, we focus on shallow social implications that manifest in the nonverbal communications. The engineering aspects are emphasized to investigate conversations using sensors ranging from audio-visual and motion sensors to those for biological and brain measurement

Figure 1: A certain amount of information is necessary even for a daily activity like setting up a presentation for a lecture.

A. The Lack of Situated Information

The advance of the information network infrastructure has connected people with each other and brought about the role of computers as a mediator in the human society. In spite of the huge amount of information made available on the net, we are still suffering from the lack of relevant information. Even for pursuing daily activities such as setting up a presentation for a lecture by connecting the PC to a projector, a certain amount of situation-specific information is needed (Figure 1). For example, the switches and controllers of presentation facilities are located in different places depending on the room, and there are subtle differences in operation sequences and semantics. Since such information is often shared by a handful of the local users and is too expensive to carefully maintain, it is often left implicit without much attention. As a result, new comers and casual users are left behind the latest updates, and disastrous failures take place from time to time. Certainly, we do not have enough situated information for daily life and need more.

B. Conversational Knowledge Process

Conversation is a handy means for people to communicate situated information. Conversation is dynamic information medium. In contrast with describing the situation in a static fashion, say by using a picture image and a written text, one can directly describe the situation by combining utterances with nonverbal communication actions, such as pointing or gaze, which are quite natural to people. For example, one might be able to communicate rather situation-dependent information as shown in Figure 2. In a more complex conversational setting with multiple participants, each participant may make structural interactions to manage shared information, as shown in Figure 3.

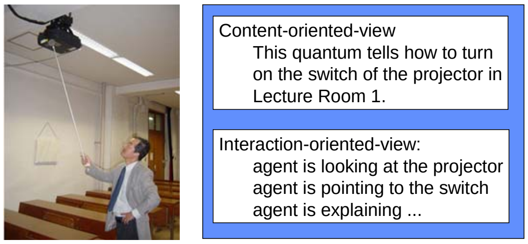
(a) Pointing and gaze (b) Posture and gaze

Figure 2: Use of conversational description style to communicate situation-dependent information.

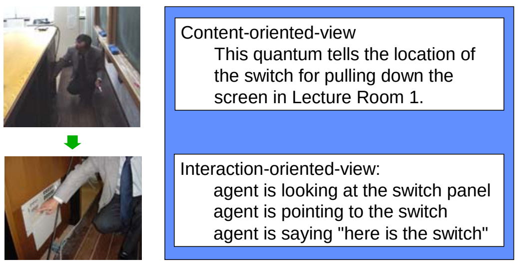
Figure 3: Multiparty conversation.

In conversation quantization [3], we introduce a conversation quantum that encapsulates interaction-oriented and content-oriented views of a conversation scene. A conversation quantum represents both content and interaction of the conversation scene. For example, a couple of conversation scenes shown in Figure 2 might be encoded in a conversational quantum as illustrated in Figure 4.

Conversational Knowledge Process is a framework for



(a) Conversation quantum for the scene in Figure 2a



(b) Conversation quantum for the scene in Figure 2b

Figure 4: Representing conversational scene by conversation quantum.

circulating conversation quanta in a community. It mainly consists of conversation quanta acquisition, accumulation and presentation, as shown in Figure 5.

Conversation quanta acquisition is a process of generating conversation quanta for a given conversation scene. So far, we have been manually encoding conversational interactions. We are now building a (semi) automated method by measuring and analyzing the participants behaviors in conversation.

Conversation quanta accumulation is a stage for accumulating conversation quanta on a server so that they can be reused in other conversation scenes. In order to allow the user to edit existing conversation quanta or to create new ones from the archive, we have developed a tool for visually manipulating the collection of conversation quanta.

Conversation quanta presentation is a stage for reproducing conversational interactions in conversation scenes. Embodied conversational agents or conversational robots are used to play a role of a participant in a conversation scene.

C. Conversation Measurement in the IMADE Room

We place much emphasis on the measuring nonverbal behaviors using the state-of-the-art sensing devices such as motion capture devices or eye trackers, rather than employing deep reasoning or planning algorithms, for we would like to gain the quality of conversation by preserving the subtle details, and also implement light-weight and robust algorithms.

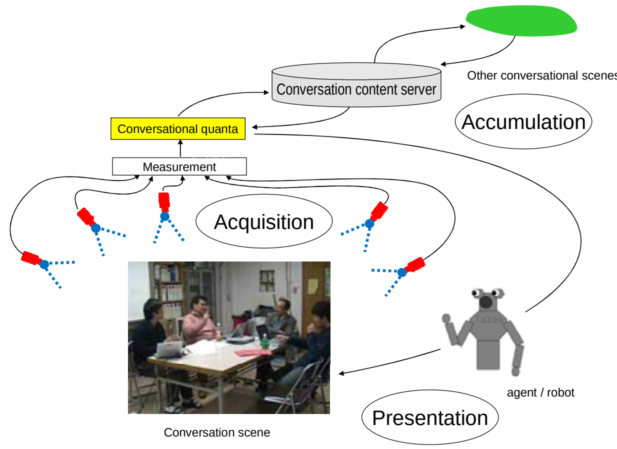
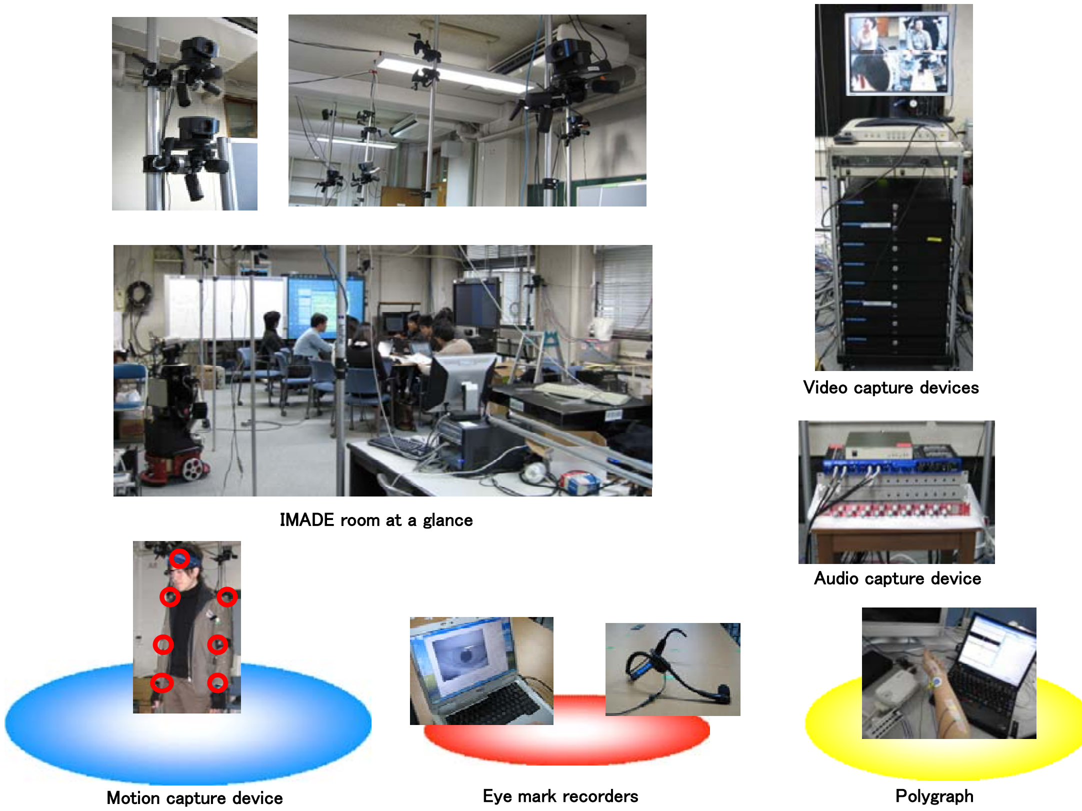


Figure 5: Conversational Knowledge Process based on conversation quantization.

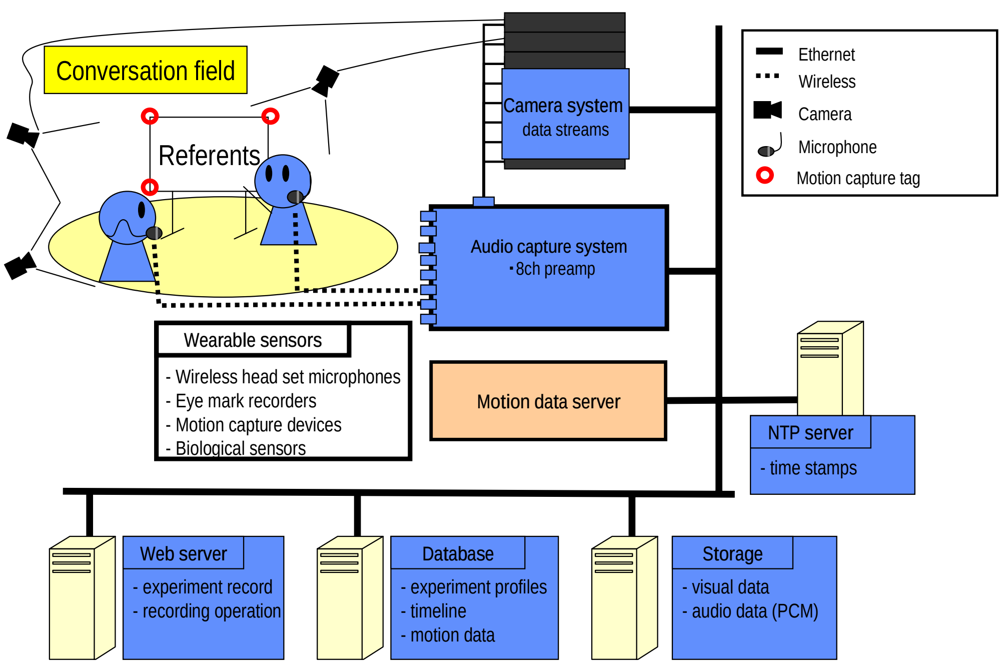
In order to study conversations by measurement and corpus building, we are developing an environment called IMADE (the real world Interaction Measurement, Analysis and Design Environment (Figure 6) [4,5]. In addition to multi-modal sensing devices, such as the wearable motion capture devices or eye mark recorders, we plan to introduce biological and brain measurement devices so that we can observe the internal activities and their interdependencies of each participant in a given conversational situation.

We have made preliminary experiments on conversation measurement and analysis. A tool called iCorpusStudio was developed for browsing, analyzing, annotating interaction corpus accumulating data obtained from experiment session [5]. In the first experiment, the behavior of group of people engaging in a collaborative design using a common display was recorded. The obtained data is being analyzed from the viewpoint of social discourse based on verbal and nonverbal interactions.

In the second experiment [5], a more complex setting was introduced to observe the dynamics of the participatory structure during the collaborative design session. In this experiment, two referents were placed in the field to see how the subject group would change the formation as discussion proceeded. Some interesting group behavior was observed that suggested the relationship between nonverbal behaviors of the participants and the group dynamics. Figure 6 demonstrates the analysis with iCorpusStudio. The analyzer is able to compare the video, audio-visual data, and annotations to study the interaction patterns observed in the session. In this example, although the subject S1 might appear to lead the migration from the left panel to the right at a glance, it turns out more likely that S1 simply dropped from the conversation and followed by the migration initiated by S2, according to the detailed analysis. Such detailed analysis is made available only by closely recording the gaze, gesture, posture, and speech of the subjects in detail, and showing an integrated view so that co-occurrences and temporal patterns of events across different modalities can be observed at a glance.



(a) Overview



(b) System configuration

Figure 5: IMADE (the real world Interaction, Measurement, Analysis and Design Environment) room [4].

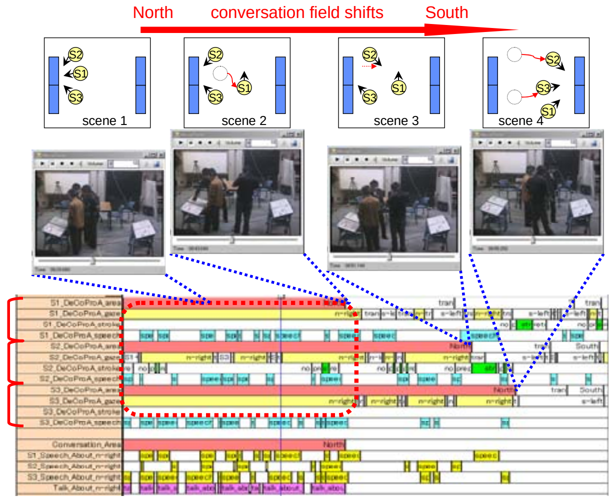


Figure 6: Analysis of multiparty interaction using iCorpusStudio [4].

Conversational Informatics

Conversational Artifacts
- Embodied conversational agents
  - Nonverbal expressions
  - Emotional aspects
  - Attentions, politeness, humors
- Conversational robots

Conversational Contents
- Conversation quantization
- Integration of natural language processing and computer vision
- Video content acquisition and editing
- Personalization

Conversational Environment Design
- Ubiquitous sensor room for recording conversation
- Real-time human proxy
- Lecture archiving system

Conversation Measurement, Analysis and Modeling
- Building and analyzing conversational corpus
- Quantitative models of nonverbal communications
- Measuring conversation atmosphere
- Theory of nonverbal interaction between humans and agents
- Model of mutual adaptation in human-robot interaction

Figure 7: Framework of Conversational Informatics [5].

D. Conversational Informatics -- State of the Art

The current development of Conversational Informatics consists of four subjects (Figure 7).

The first subject is conversational artifacts (embodied conversational agents or conversational robots) that can participate in human conversations. Our study involves algorithms for interpreting and presenting nonverbal expressions to permit the user to interact with them in a conversational fashion, not only with natural language but also with eye gaze, facial expressions, gestures, or other nonverbal communication means. Socio-emotional implications of conversation such as attentions, politeness, friendliness, or personality are investigated with great interest.

The second subject is about manipulating conversational contents that encapsulate information arising in conversation scenes. Techniques are being developed for accumulating, editing, and converting conversational contents, using natural language processing, computer vision, and human computer interaction.

The third subject is conversation environment design. The primary goal is designing an intelligent environment that can sense and augment the conversational interactions. Work is in progress to provide situated information supports by combining wearable or environment sensors and displays in conversation scenes ranging from poster sessions to large classrooms.

The last subject is conversation measurement, analysis and modeling, driven by scientific interest. Introduction of powerful sensing technologies significantly accelerates the study. It will not only permit a data-driven quantitative understanding of conversational behaviors but also enable a corpus-based development of conversational systems that are more robust and sophisticated than those by pure programming.

In the next three sections, I would like to survey recent developments in Conversational Informatics.

III. Conversational Artifacts

The role of conversational artifacts is to provide a conversational interface with the user. Conversational artifacts consist of two major categories depending on whether they have a physical body or not.

A. Embodied Conversational Agents

Embodied conversational agents are interactive synthetic characters that have CG-based embodiment. Embodied conversational agents have a relatively long history of development. They originate from synthetic characters and natural language dialogue systems. More sophisticated nonverbal interaction functions have been incorporated as the technologies advance [6, 7]. The generic, component-based platform is being employed, rather than hard-wired application-specific architecture. More emphasis is placed on simulating subtle features of nonverbal expressions based on corpus and allowing large-scale rich content to be referred to in conversation.

The GECA (Generic ECA) is a generic framework for building an ECA system on multiple servers connected with each other by a computer network [8]. GECA allows for mediating and transporting data stream and command messages among software modules. It provides with a high-level protocol for exchanging XML messages among components such as input sensors, inference engines, the emotion model, the personality model, the dialogue manager, the face and body animation engines, etc. An application programming interface is made available on main-stream operating systems so that the programmer can easily adapt ECA software modules to incorporate into the GECA platform. The blackboard model is employed as the backbone.

GECA has been implemented and applied for various applications involving a navigation agent, a quiz agent, and a pedagogical agent for teaching cross cultural communication.

A navigation agent was designed to make a spatial navigation for the user. The user can talk to the navigation agent about objects in the background, by combining natural language, hand pointing and head movements. In response, the navigation agent combines speech with eye gaze, facial expressions, hand gestures, and postures to guide the user in a simulated place, possibly by moving around there.

The quiz agent was designed to entertain visitors at open house events of a research institute [9]. An interactive synthetic character is displayed on a large screen. When a visitor arrives, the quiz agent will give the user a number of puzzles. Each puzzle consists of a question followed by several alternatives. When the visitor chooses one, the quiz agent will tell whether the choice is correct or not, and explain the correct answer when the visitor's answer is wrong. Touch panel was chosen as the input device for complex sensors were considered to be unstable and might worry the visitor. Emotional feedback was implemented, using the PAD space model. Positive stimulus will be given to the emotion and mood when the visitor tries to answer the quiz. Even higher values will be given if the answer is correct. In contrast, negative stimulus will be given when the answer is wrong. The value on boredom axis will be increased when no input is given from the visitor in a certain amount of time.

This quiz agent was demonstrated to the public in a one-day public open lab event of NFRI (National Food Research Institute) on April 20th, 2007. In the demonstration session lasting for six hours, 307 visitors in small groups played the kiosk and 87 game sessions were run. The analysis of questionnaire revealed that most of the visitors enjoyed the game and felt that the knowledge explained by the agent was trustable.

B. Conversational Robots

Robots' physical embodiment normally yields a high presence and strong social implication in communication. Efforts have been made to build conversational robots that can participate in conversations. In early days of development, communication was made only with speech interface. Recent implementations, in contrast, place much emphasis on the nonverbal communication abilities. Nishida et al [10] has proposed the notion of robot as an embodied knowledge medium, where robots bear a role of mediating knowledge among people. The listener and presenter robots were prototyped to investigate the feasibility of the idea. The listener robot was designed to videotape critical scenes while interacting with an instructor. In the meanwhile, the presenter robot was designed to assist a novice user by showing appropriate video clip on a small display attached on the arm.

Both the listener and presenter robots were designed to detect critical behaviors of the user such as gaze or pointing to coordinate behaviors by intentionally making communicative acts such joint attention.

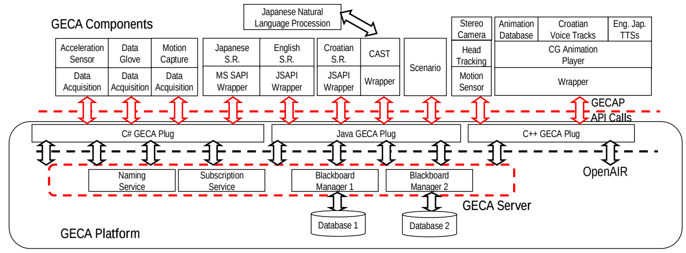


Figure 8: The Architecture of GECA [8].

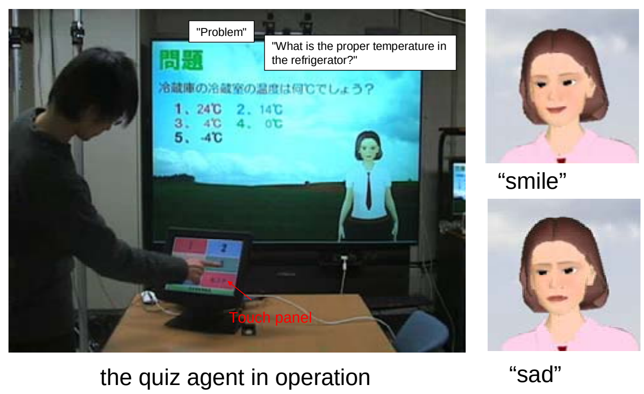


Figure 9: The quiz agent [9].

In case of the listener robot, for example, ten 3D position sensors were attached to the instructor's body, and several 3D position sensors were used to identify the location of the salient objects in the environment. The user's status is sensed by a motion capture device and interpreted using Bayesian networks.

As a result, the listener robot can distinguish transitions of critical conversation modes, such as the talking-to, or talking-about modes.

Figure 10 shows how the listener robot interacts with the human user. In Figure 10a, the listener robot makes a joint attention according to the instructor’s pointing gesture. Figure 10b shows the image of the object captured by the listener robot's eyes at that moment. Figure 10c and d shows how the listener robot interacts with two instructors. In Figure 10c, the two instructors are talking to the robot, and the robot replies to the person in the left, while in Figure 10d, both the instructor in the left and the robot are looking at the work of the instructor in the right.

Figure 11 illustrates the way the presenter robot behaves. The presenter robot coordinates eye gaze, posture, and motion

(a) Joint attention

(b) Image from the robot's eye

(c) *A* is talking

(d) Robot is looking at *B*

Figure 10: The listener robot.

according to the user's behavior. As a number of experimental evaluations, it turned out that the user was able to complete the task about 63% of the baseline setting with the fixed display position, with less (about 50%) interaction time and less frequency (54%) of interaction.

IV. CONVERSATIONAL CONTENT

Technologies are being developed to help people create and manage a large amount of contents collected from conversations.

A. *Visual Accumulation of Conversational Contents*

The Sustainable Knowledge Globe (SKG) [11] is a system that allows the user to visually accumulate a large amount of conversational contents on the sphere surface so that s/he can a long-term relationship with them to complement the limitation of her/his biological memory. Conversational contents may be grouped into a tree structure so that the user can manipulate them as a group. A graphical user interface is employed to continuously zoom in/out the any region of the sphere surface, as shown in Figure 12. A linear zooming method is employed to avoid distortion of the landscape on the sphere surface. In order to help the user visually recognize the tree structure, we have introduced nesting contours. Embodied conversational agent was installed on the SKG system to navigate the user by presenting the content interactively.

B. *Media Conversion*

Media conversion is a powerful method to obtain conversational contents from a huge amount of legacy contents, such as natural language documents or archived videos. Kurohashi et al [12] developed a method for automatically converting a collection of series of short documents called knowledge cards into conversational contents consisting of spoken language scenario and summarization slides that can be

Figure 11: The presenter robot [9].

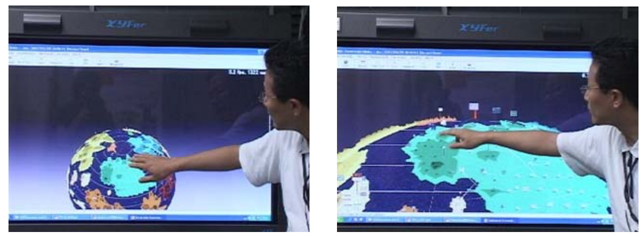


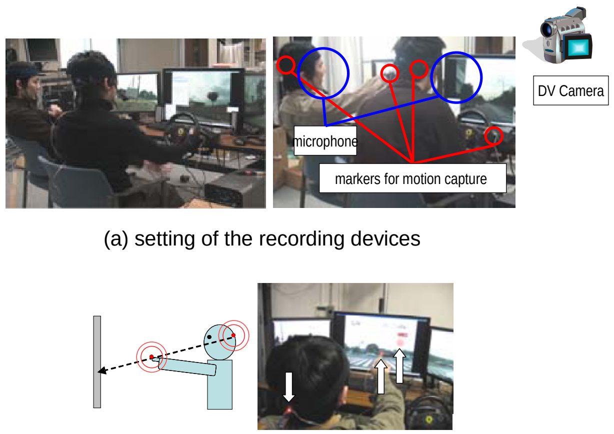
Figure 12: Sustainable Knowledge Globe [11].

automatically presented in a conversational fashion using embodied conversational agents. The method is based on corpus-driven natural language processing techniques for automatic construction of large-scale case frame, analysis of predicate-argument structure, and discourse structure analysis.

Nakamura [13] proposed an automated video content acquisition and editing for small meetings. The system follows two stages to generate a video summary for a dialogue. The first stage is content capture. The environment cameras and content production cameras are used for video capture. The system controls cameras to keep typical picture compositions such as close-up/bust shot, over-the-shoulder shot, or long shots. Contents capturing camera modules detect and track the face of the participants. The second stage is editing conversational scenes. At this stage, video streams from contents capturing cameras are edited into one stream, by maximally satisfying constraints extracted from various camera switching techniques. Speeches, motion, facial expressions, object movements, etc are taken into account for the processing.

V. CONVERSATIONAL ENVIRONMENT DESIGN

The goal of conversational environment design is to provide a smart environment that allows people to pursue effective knowledge creation through conversations. Approaches vary depending on the size of the conversation environment, whether the environment is distributed or not, how much auxiliary devices can be introduced, how much quality is required, how much cooperation is expected from the participants, how much cost can be spent on the environment,



(a) setting of the recording devices

(b) detecting pointing gestures

Figure 13: An augmented conversational environment for a driving simulator [14].

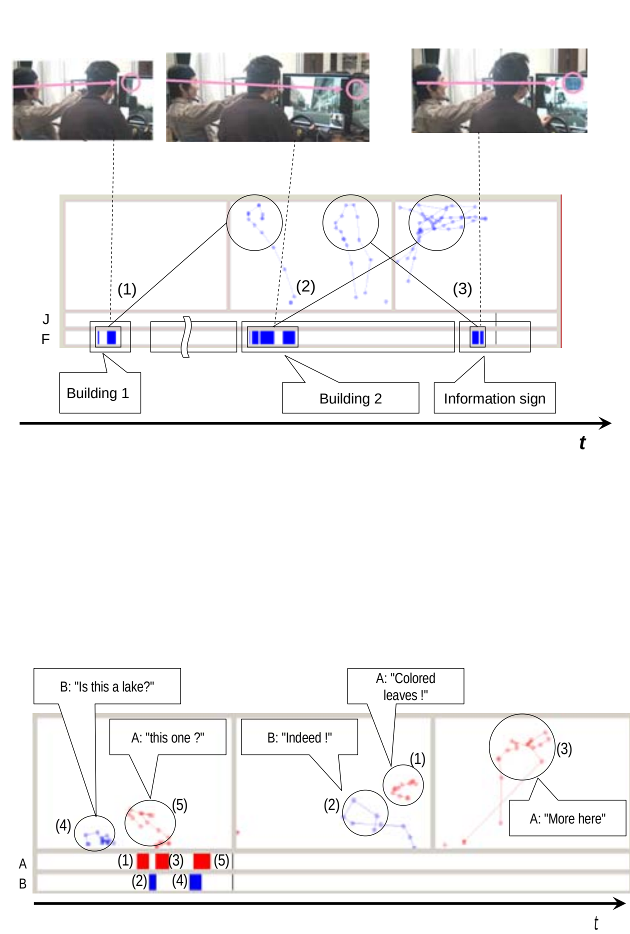


Figure 14: Pointing gestures and conversation discourse observed in the setting shown in Figure 13 [14].

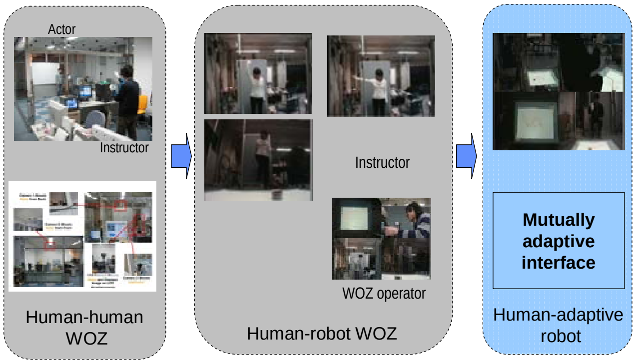


Figure 15: Three stage approach to mutual adaptation [28].

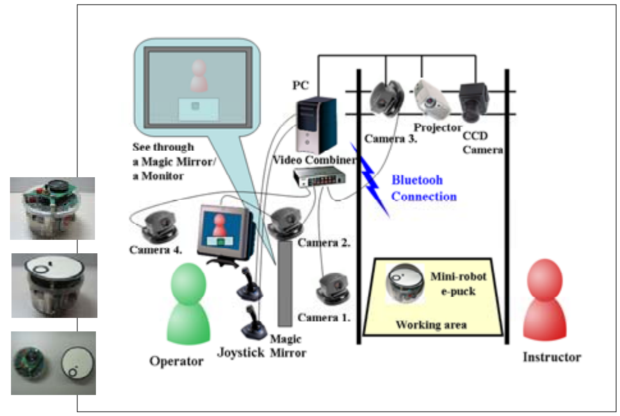


Figure 16: Facilities for measuring mutual adaptation in a human-robot WOZ [28].

speaker to gaze at the current speaker toward the end of his turn, but also a tendency for the other listener, who keeps silent during the next turn, to gaze at the next speaker around the end of the current turn and before the next speaker starts speaking." Ueda and Ohmoto prototyped a real-time system that can discriminate lies by measuring gaze directions and facial feature points [22]. The system can measure gaze directions and facial feature points, while allowing the user to move head position and orientation during the measurement and without requesting the user to place one or more markers on the face or preparing a face model in advance. Nagaoka et al [23] made a comprehensive survey of embodied synchrony (phenomenon of synchronization or similarity of nonverbal behaviors among participants) reported in diverse literature. The embodied synchrony manifests as body movement and gestures, facial behavior, vocal behavior, or physiological reactions. They surveyed the measurement and quantification techniques that have been employed in previous studies. They also attempted to attribute the embodied synchrony to interpersonal relations. Rutkowski and Mandic [24] addressed characterization of what may be called a communication atmosphere. They proposed the communication atmosphere space consisting of three dimensions: environmental, communicative, and emotional. They used audio-visual signal tracking to show the measurement for a handful of example data.

Mohammad and Nishida study human-robot communication of intentions using nonverbal behaviors. Early results include the use of interactive perception to establish and maintain joint intention [25] and a social robot that can express its internal state and intention to humans in a natural way using nonverbal feedback [26].

Mutual adaptation is a phenomenon we believe to exist between multiple learning agents being adapting with each other. Xu et al [27, 28] study mutual adaptation by taking a three stage approach consisting of a human-human WOZ experiment, a human-robot WOZ experiment, and a human-adaptive robot experiment (Figure 15). Instead of directly diving into the third stage, we observe in detail how

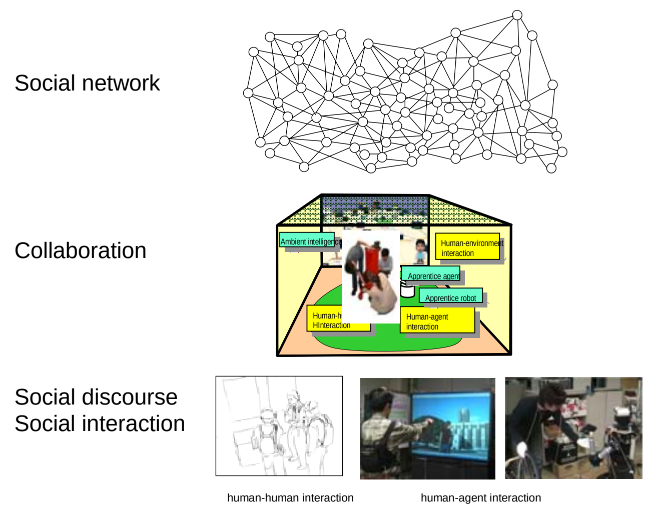


Figure 17: Three layer mode of social intelligence.

people adapt with each other and how people improve the protocols for interacting with robots. Figure 16 shows the experimental environment we developed for measuring mutual adaptation in a human-robot WOZ. We use a small mobile robot controlled by a hidden operator as if the robot was autonomous. By observing how the instructor interacts with the operator, we try to solicit the detailed observation of mutual adaptation.

VII. Social Intelligence Design

Social Intelligence Design [29] aimed at the understanding and augmentation of social intelligence for collective problem solving and learning. Social intelligence may manifest at the three levels (Figure 17). The base level comprises quick interactions at the milliseconds order where social intelligence is used to establish basic communications. The medium level encompasses a collaboration or negotiation in a small group to coordinate joint actions. The top level manifests at the community level to integrate individual intelligences into a collective one. Conversational Informatics discussed in this article is most relevant to the social discourse level. The upper levels may have closer relationship with Web Intelligence. In order to realize Human-Centered Web Intelligence, we need to study how the layers interact with each other.

REFERENCES

[1] N. Zhong, J. Liu, Y.Y. Yao, "Envisioning Intelligent Information Technologies (iIT) from the Stand-Point of Web Intelligence (WI)," *Communications of the ACM*, 50(3), 2007, 89-94.

[2] T. Nishida, *Conversational Informatics: an Engineering Approach*. London: John Wiley & Sons Ltd, in press.

[3] T. Nishida, "Conversation Quantization for Conversational Knowledge Process," S. Bhalla (Ed.): *DNIS 2005*, LNCS 3433, Springer, 2005, pp. 15 – 33.

[4] S. Sumi, M. Bono, H. Kijima, and T. Nishida, "The IMADE (real world Interaction Measurement, Analysis and Design Environment)," unpublished.

[5] H. Kijima, M. Bono, Y. Sumi, and T. Nishida, "Development of the Environment for Multimodal Interaction Analysis," to be presented at SIG-HCI, IPSJ, Japan, 2007.

[6] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill (eds.), *Embodied Conversational Agents*, MIT Press, 2000.

[7] H. Prendinger and M. Ishizuka M. (eds.), *Life-Like Characters: Tools, Affective Functions, and Applications*, Springer, 2004.

[8] H. Huang, A. Cerekovic, I. Pandzic, Y. Nakano, and T. Nishida, "Scripting Human-Agent Interactions in a Generic ECA Framework," to be presented at AI-2007, 2007, Cambridge, UK.

[9] H. Huang, T. Inoue, A. Cerekovic, I. Pandzic, Y. Nakano, and T. Nishida, "A Quiz Game Console Based on a Generic Embodied Conversational Agent Framework," *Proceedings of IVA 2007*, pp. 383-384.

[10] T. Nishida, K. Terada, T. Tajima, M. Hatakeyama, Y. Ogasawara, Y. Sumi, Y. Xu, Y. Mohammad, K. Tarasenko, T. Ohya, and T. Hiramatsu: Towards Robots as an Embodied Knowledge Medium, Invited Paper, Special Section on Human Communication II, *IEICE TRANSACTIONS on Information and Systems*, Vol. E89-D, No. 6, 2006, pp. 1768-1780.

[11] K. Kubota, Y. Sumi, and T. Nishida, "Conversation Quantization and Sustainable Knowledge Globe," Chapter 10 of [2].

[12] S. Kurohashi, D. Kawahara, N. Kaji, and T. Shibata, "Automatic Text Presentation for the Conversational Knowledge Process," Chapter 11 of [2].

[13] Y. Nakamura, "Video Content Acquisition and Editing for Conversation Scenes," Chapter 12 of [2].

[14] G. Okamura, H. Kubota, Y. Sumi, T. Nishida, H. Tsukahara, and H. Iwasaki, "Quantization and Reuse of Driving Conversations," Journal of IPSJ (in Japanese, to appear).

[15] L. Merckel, and T. Nishida, "Solution of the Perspective-Three-Point Problem," in *Proceedings IEA/AIE 2007*, pp. 324-333.

[16] H. Kubota, M. Takahashi, K. Satoh, Y. Kawaguchi, S. Nomura, Y. Sumi, and T. Nishida, "Conversation Quantization for Informal Information Circulation in a Community, The Fourth International Workshop on Social Intelligence Design (SID 2005), Stanford, USA.

[17] K. Saito, H. Kubota, Y. Sumi, and T. Nishida, "Analysis of Conversation Quanta for Conversational Knowledge Circulations," *Journal of Universal Computer Science*, vol. 13, no. 2, 2007, pp. 177-185.

[18] Y. Sumi, K. Mase, and T. Nishida, "Conversational Content Acquisition by Ubiquitous Sensors," Chapter 14 of [2].

[19] R. Taniguchi and D. Arita, "Real-time Human Proxy," Chapter 15 of [2].

[20] S. Nishiguchi, K. Kakusho, and M. Minoh, "Lecture Archiving System" Chapter 16 of [2].

[21] Y. Den and M. Enomoto, "A Scientific Approach to Conversational Informatics: Description, Analysis, and Modeling of Human Conversation," Chapter 17 of [2].

[22] Y. Ohmoto, K. Ueda, and T. Ohno, "Real-time system for measuring gaze direction and facial features: Towards automatic discrimination of lies using diverse nonverbal information," *AI & Society*, (online first).

[23] C. Nagaoka, M. Komori, and S. Yoshikawa, "Embodied Synchrony in Conversation," Chapter 18 of [2].

[24] T. Rutkowski and D. Mandic, "Modeling Communication Atmosphere," Chapter 19 of [2].

[25] Y. Mohammad and T. Nishida, "NaturalDraw: interactive perception based drawing for everyone," *Intelligent User Interfaces* 2007, pp. 251-260

[26] Y. Mohammad and T. Nishida, "TalkBack: Feedback From a Miniature Robot," to be presented at Australian AI Conference.

[27] Y. Xu, K. Ueda, T. Komatsu, T. Okadome, T. Hattori, Y. Sumi, and T. Nishida, "WOZ Experiments for Understanding Mutual Adaptation," AI&Society, online first.

[28] Y. Xu, M. Guillemot and T.iNishida, "An experiment study of gesture-based human-robot interface," IEEE/ICME International Conference on Complex Medical Engineering-CME2007, Beijing, China, 2007, pp. 458-464.

[29] T. Nishida, "Social Intelligence Design and Human Computing," *Artificial Intelligence for Human Computing 2007*, pp. 190-214.

Toyoaki Nishida (M'01) was born in 1954 in Kyoto, Japan. He is a professor of Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University. He received the Doctor of Engineering degree from Kyoto University in 1984. He received the B.E., the M.E., and the Doctor of Engineering degrees from Kyoto University in 1977, 1979, and 1984, respectively.

His research centers on artificial intelligence and human computer interaction. His current research focuses on social intelligence design and communicative intelligence. In 2001, he founded a series of international workshops on social intelligence design (see http://www.ii.ist.i.kyoto-u.ac.jp/sid/ for more details). Then, he broadened the scope of research to include understanding and augmenting conversational communication, and opened up a new field of research called Conversational Informatics. Currently, he leads several projects on social intelligence design and conversational informatics.

Prof. Nishida is a member of the board of directors of IPS (Information Processing Society) of Japan and JSAI (Japanese Society for Artificial Intelligence). He serves as an editorial board member of several academic journals, including Web Intelligence and Agent Systems, AI & Society, and Journal of JSAI (editor-in-chief).