Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Providing Adaptive Support in Computer Supported Collaboration Environments

Kinshuk[1], Yoshitaka Sakurai[2], Kohei Takada[2], Sabine Graf[1], Ardah Zarypolla[2], Setsuo Tsuruta[2]

[1] School of Computing and Information Systems
Athabasca University
Alberta, Canada
kinshuk@athabascau.ca, Sabine.Graf@ieee.org

[2]School of Information Environment
Tokyo Denki University
Chiba, Japan
{ysakurai, tsuruta}@sie.dendai.ac.jp

***Abstract*— Many misunderstandings can occur during remote interaction due to different user domain competency levels, different cognitive capacity of users as well as different user backgrounds. In this paper, we propose an adaptive keyword/summary presentation approach that aims at identifying potential misunderstandings of individual users and provide these users with effective and personalized content of the current discussion. Our approach is developed for virtual worlds and tested and implemented based on the Wonderland Project. In order to evaluate our approach, a practical scenario has been designed and tested, which demonstrates how the system enriches the cyberspace for collaboration by making adaptive use of keyword/summary presentation.**

Keywords*—. *Adaptive, keyword/summary, misunderstandings, cyberspace, collaboration

I. Introduction

People physically close to each other can better understand each other and easily keep track of the status and activities of others. On the other hand, in remote meetings more misunderstandings can happen due to different user backgrounds, different competency levels, and different personal agendas. Furthermore, remote connections mean participants are in their own environment with all its nuances rather than with other participants in the same room, which means the chances of external disturbance are higher.

In this paper, we introduce an adaptive keyword/ summary presentation feature for enriching remote collaborations in cyberspaces. This approach gathers various information about the users such as a cognitive profile, biological sensor information, data from webcam and acceleration meter, and history data from interactions via mouse, keyboard and microphone. Based on this information, the situation of the user is assessed and potential misunderstandings are identified. The adaptive keyword/summary presentation feature provides users with personalized support for individually overcoming or avoiding potential misunderstandings, enabling the users to understand the current discussion easily, clearing up the occurred misunderstandings, and therefore making the remote collaboration a success.

Current research works in this area focus on providing a wide range of personal assistance rather than considering the users' individual situation. For example, Enembreck and Barthes [1] developed a personal assistant to improve collaborations among group members, which provides intelligent access to documents, communication tools and domain tasks, fosters group awareness, and helps sharing information among team members. Furthermore, researchers have focused their attention on the information retrieval speed. For example, Sia, Cho, and Cho [2] exploited the RSS aggregation services [3, 4] that monitor the data sources to retrieve new content quickly and provide its subscribers with fast news alerts.

Meanwhile, Sakurai et al. [5] developed a system that augments the cyberspace by using multiple biological sensor information as the basis for a warning feature to support the reliable communication in a remote cooperative environment and provide personal warning/stressing assistance. In one of our previous research works [6], we proposed an "Enriched Cyberspace" approach including various features for dependable web based computer supported collaborative learning (CSCL) to overcome various misunderstanding problems in remote learning situations.

This paper investigates individual factors which can results in various individual misunderstandings and proposes an adaptive summary/keyword presentation feature which aims at providing individual support for users in remote collaborations based on their individual situation. In the next section, a system architecture for such adaptive support in remote collaboration is described. Section 3 presents the investigated types of misunderstandings, the adaptive features which can help in preventing/sorting out these misunderstandings and how such misunderstandings can be identified. Section 4 presents the scenario based evaluation of the adaptive feature and Section 5 concludes the paper.

II. System Architecture

The adaptive keyword/summary presentation feature is designed for supporting remote collaborations in cyberspace environments. In this research, the Project Wonderland [7] was used as the basis. The Project Wonderland, developed by Sun Labs, provides a 3D collaborative virtual environment that lets users' avatars interact with shared artifacts, communicate with other avatars, and share live applications such as web browsers. Furthermore, Julius [8], a language-independent, highperformance speech recognition software was used. Julius has the word accuracy of 95%, and over 90% in real-time processing [8]. A grammar-based recognizer called "Julian" is integrated in Julius to recognize speech data and convert it to text data.

Julius opens the specified input audio stream and enters a recognition loop which can be used to support users in their collaborations where misunderstandings can occur. In this loop, Julius detects speech segments from the microphone input stream and processes the data. The output results can be shown on the speaker's chat frame and distributed to individual users who are identified as having certain type of misunderstanding.

In order to support users in collaborating with each other and reduce misunderstandings in their collaboration, Julius' original components, namely the keyword vocabulary file and the grammar modules file are modified. Our extensions aimed at enabling Julius to extract keywords and summaries from speaker's lecture and distribute them via the wonderland server to those wonderland clients, for whom possible misunderstandings have been identified. This process is depicted in Fig. 1.

Our keyword/summary presentation approach for preventing/dealing with misunderstandings in web-based collaborative systems is evaluated by enriching wonderland cyberspace environment. Enriching wonderland cyberspace is based on identifying and assessing the situations of remote users through fusing information of multiple biological sensors and the related user profile contexts. Fig.2. shows this concept in more detail.

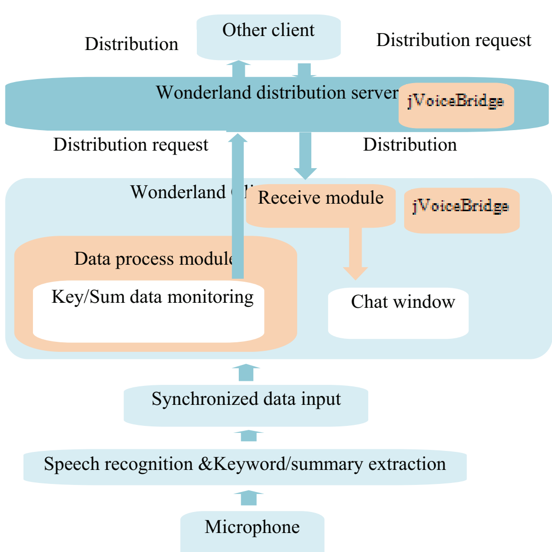

Figure 1. Implementation process of keyword/ summary extraction and presentation

Various real-time input parameters obtained from multi-biometric sensors and individual user profiles are used to obtain information about the users' situation during the collaborative activities. The sensors include a sensor mouse, web camera, accelerometer, and microphone. The sensor mouse provides biological information about skin temperature and blood pulse frequency. The web camera is used for facial expression recognition (e.g., identifying whether a user is yawning) and provide information about head movements which can in turn provide information about the user's level of distraction (e.g., if sidewise head movement is identified, the user might be looking somewhere else or talking with another person face-to-face and is therefore distracted from the collaborative online activities). The accelerometer gathers data about the chair movements and the voice signal from the microphone is used for speech recognition.

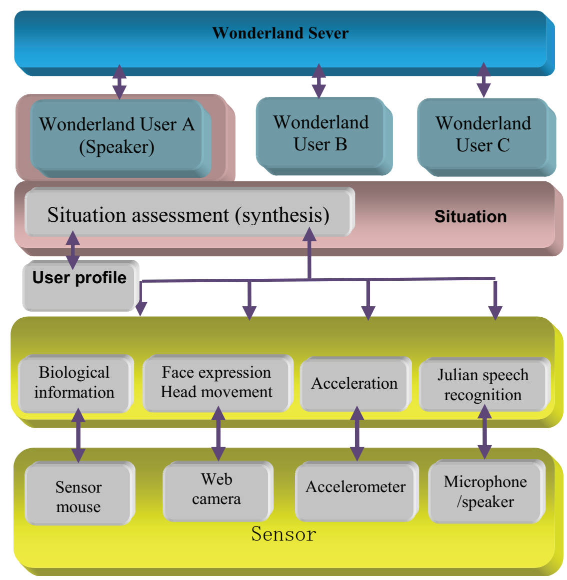

Figure 2. Enriching Wonderland cyberspace

All these data are archived in the user profile. Furthermore, the user profile includes background information about the users such as their location information history, preference for medium of message, learning styles, domain competence, users' areas of interest, the users' cognitive profile including their short term memory capacity, inductive reasoning ability, associative learning skills and information processing speed, the current speaker's assessment about individual users' understanding level during the remote interaction, and a record of users' interactions via keyboard and mouse. Of special interest for the currently investigated types of misunderstandings are the users' short term memory capacity and associative learning skill from cognitive profile, users' domain competency, users' interactions via keyboard and mouse which give indications about the distractions

Table 1. Misunderstandings due to personal situation

Type	Description	Indication	Cause	Input parameters	Solution
Information decay	Forgetting previous information	• Checking the past interaction history • Asking clarifications • Trying to read previous notes	• Low short term memory	• Microphone (voice signal) • Cognitive profile (via interaction pattern) • User interaction (KB/mouse action on non-related window)	Keyword presentation
Misapplying knowledge	Misconceptions: Problems in understanding the essence of current discussion	• Asking question(to speaker) • Yawn	• Low Associative learning skill • Tiredness(Bore) • Difference of domain competence	• Webcam • Microphone (voice signal) • Cognitive profile (via interaction pattern) • Blood pulse frequency	Summary presentation
Not hearing	Missing completely other speaker's talk	• Sidewise look • Sleepy	• External disturbance • Tiredness(Fatigue)	• Mouse sensor (Blood pulse frequency, Skin temperature) • Webcam(Up/down and Sidewise head movement) • Acceleration meter (Chair movement) • Microphone (voice signal)	Keyword/ Summary presentation

experienced by the users during a remote session, and various biometric data gathered from the sensors.

As can be seen from the description above, various sources are used as input parameters for the identification and assessment of the situation of the users. This information is used to validate the reliability of analysis of misunderstandings and is synthesized by a situation synthesis program. The result of this synthesis is shown on the speaker's display and appropriate adaptive keywords/summary presentation is displayed to individual users.

The situation synthesis program is included in *Fusion Engine* that imitates human intelligence for situation assessment through diverse information pieces such as multiple sensor information. For reliable situation assessment, two kinds of knowledge are interpreted: *Information synthesis knowledge* synthesizes information from input parameters to further synthesize information related to symptoms (e.g., concentration and flashing in the brain). Next, information on situations is synthesized from these symptoms. *Information analysis knowledge* analyzes synthesized symptoms and situations, which can then be used by the system to predict and acquire information necessary for solving contradictions, interpolating the lost information, or just confirmation.

III. Adaptive Features for dealing with Misunderstandings

This research is focused on identifying misunderstandings due to personal situations. Table 1 shows the investigated types of misunderstandings, brief descriptions, typical indications for such misunderstandings, their causes, the necessary input parameters for identifying these misunderstanding, and solutions to resolve them.

In this paper, we focus on three types of misunderstandings: *Information decay* means users easily forget previously discussed information, *misapplying knowledge* means users have problems in understanding the essence of the current discussion, and *not hearing* means users completely miss the others' talk.

For these three types of misunderstandings, we propose two adaptive features, namely the keyword presentation feature and the summary presentation feature, which aim at supporting users during remote collaboration and reducing misunderstandings.

The keyword presentation feature presents keywords of current discussion on the user's display when the system identifies user's potential misunderstanding problems from information decays. The system checks voice interaction from user's microphone, user's interaction from the keyboard/mouse actions on non-related windows (e.g., trying to read previous notes), and the user's cognitive profile with respect to the user's short term memory capacity. If the user has voice interaction with speaker through microphone, or if the user reads previous notes according to keyboard/mouse sensor data and the user has low short term memory capacity, then system identifies this as an indication for potential misunderstandings with respect to information decay. In such scenario, the system presents keywords to the user.

The summary presentation feature provides users with a summary of the current discussion if the system identifies problems in user's understanding about the essence of the current discussion. The system checks voice interaction from user's microphone in order to find out whether the user asks questions to the speaker, the web camera in order to identify if the user is distracted (e.g., looking to somebody else) or if the user is tired (e.g., is often yawning) which can result in a lack of concentration, the blood pulse frequency from the mouse sensors which also can give indications about the user's tiredness, the user's domain competence, and the associative learning skills from the cognitive profile. If the user has voice interaction with speaker through microphone, if yawn/tiredness is detected by webcam and mouse sensor, if the system identifies lack of domain competence, and if associative learning skills are low according to the cognitive profile, the potential for a misunderstanding with respect to misapplying knowledge is identified as high and the system presents a summary of the speaker's conversation in order to prevent such a misunderstanding. Furthermore, summary presentation is provided, if the user is missing parts of the explanation/talk of the speaker due to external disturbance which can be detected from the webcam (e.g., sidewise movement).

Both summary and keyword presentation features are provided if the user is not hearing the speaker's talk and missing it completely. Therefore, the system aims at detecting the user's sleepy state by various data from various sensors and the degree of external disturbance such as sidewise looks by the user. In order to get this information, data about blood pulse frequency and skin temperature are gathered from the sensor mouse. Furthermore, the webcam, acceleration meter, and the microphone provide additional data. If the user has no voice interaction with the speaker through the microphone, if no up/down and sidewise head movement is detected by webcam and user's sleepy state is detected due to his/her low blood pulse frequency gathered from the mouse sensor, the system activates individual prompt feature by using pop-up windows/voice prompting, and then activates the keyword and summary presentation to help the user to keep up with the speaker's explanation.

IV. Evaluation

The adaptive keyword/summary presentation features are evaluated by a designed scenario in Wonderland enriching cyberspace in order to demonstrate its feasibility. One speaker, identified as A, and three listeners B, C, and D are attending an online collaborative team session and follow content which is taught by speaker A.

A. *Evaluation Scenario*

In order to evaluate the adaptive keyword/summary presentation features, a scenario dealing with the fallacies of distributed computing is used. The fallacies of distributed computing are a set of common but flawed assumptions made by programmers when first developing distributed applications [9]. The fallacies are summarized as follows:

1. The network is reliable.
2. Latency is zero.
3. Bandwidth is infinite.
4. The network is secure.
5. Topology does not change.
6. There is one administrator.
7. Transport cost is zero.
8. The network is homogeneous.
9. System clocks are identical.

In our scenario, user D has low short term memory capacity and therefore might easily forget parts of the speaker's talk. Furthermore, user C is tired and therefore might easily misunderstand the speaker's explanations. Finally, user B did not get enough sleep at the previous night so that he is sleepy at the meeting.

B. *Evaluation findings*

Simulating the abovementioned scenario, the system solves the users' various misunderstandings by considering indications of potential misunderstanding situations which result from users' various sensor data and user profile information. In the following paragraphs, we describe how the system detects potential misunderstandings and provides users with adaptive features in order to overcome possible misunderstandings.

Test Case 1: First, we take a look at user D, who is at risk of forgetting parts of the speaker's explanations. The system has access to data about microphone signal through user's microphone, keyboard/mouse actions of user D through tracking of interaction patterns, and information about user D's cognitive profile. The system identifies that user D has voice interaction signal with speaker A, and user D's mouse is moving in order to try to read previous notes. Furthermore, user D has low short term memory, which has been detected from the user profile. These are identified by the system as indications about information decay for user D (Fig. 3). Therefore, the system activates the keyword presentation feature in order to help user D keep up with speaker explanation and remember forgotten parts (Fig. 4). A formal presentation of this test case can be seen in Figure 5.

Figure 3. User D with low short term memory

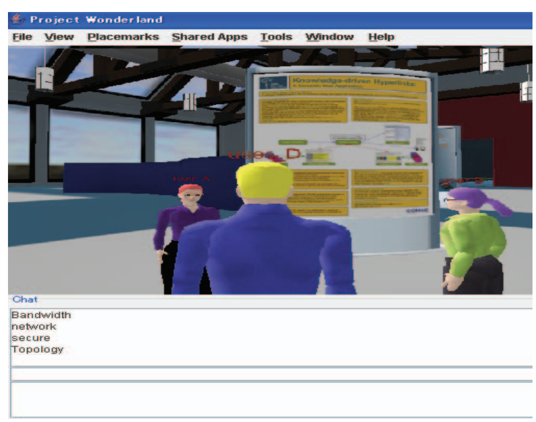

Figure 4. Keyword was sent to user D

[Test Case 1]: Information decay

Problem: User D forgets previous explanation given by speaker A.

System analysis:

- Cognitive profile
- Keyboard/mouse actions
- Voice interaction

System action:

If user D has low short term memory, is trying to read previous information, and has voice signal interaction with speaker

{Display keywords of speaker's talk that was forgotten}

Adaptive support:

User D can keep up with speaker's explanation.

Figure 5. Formal presentation of test case 1

Test Case 2: User C is bored/tired and therefore might easily misunderstand speaker A's explanation. The system has access to data about microphone signal through user's microphone, the user C's face expression through the webcam, blood pulse frequency from mouse sensor, and cognitive information from the user C's cognitive profile. User C has low associative learning skills. Additionally, the system identified that there are voice signals from user C's microphone, indicating that user C is asking questions to the speaker. Furthermore, user C is yawning and his blood pulse frequency and skin temperature is decreasing, indicating bore/tiredness. Based on these indications, the system considers user C to have problems in understanding the essence of the current discussion. Therefore, the system activates summary presentation to help user C to keep up with the current discussion (Fig. 6). A formal presentation of this test case can be seen in Figure 7.

Figure 6. Summary is displayed for user C with low associative learning skill

[Test Case 2]: Misapplying knowledge

Problem: User C has problems in understanding the essence of the current discussion

System analysis:

- Webcam: Yawn
- Mouse sensor: Blood pulse frequency and skin temperature of user C
- Cognitive profile
- Voice interaction

System action:

If user C has low associative learning skills, a voice signal interaction with the speaker, and is yawning often

{Display summary of speaker's talk that was not understood}

Adaptive support:

User C can keep up with speaker's explanation.

Figure 7. Formal presentation of test case 2

Test Case 3: User B is listening to the speaker's explanation (Fig. 8). However, user B is sleepy and therefore, after some minutes, he is not hearing the explanation given by the speaker any more (Fig. 9). The system has access to data about the user B's face direction through the webcam (up/down and sidewise head movement), chair movement from acceleration meter, and the skin temperature and blood pulse frequency through mouse sensors. The system identifies that user B's skin temperature and blood pulse frequency is decreasing. Furthermore, there are no chair movements and user B's head is moving up and down (Fig. 9). Based on these indications, the system considers user B as sleepy and makes an alert sound. Then the system activates the keyword/summary presentation features in order to help the user to keep up with the speaker's explanation (Fig. 10). A formal presentation of this test case can be seen in Figure 11.

Figure 8. User A is speaking

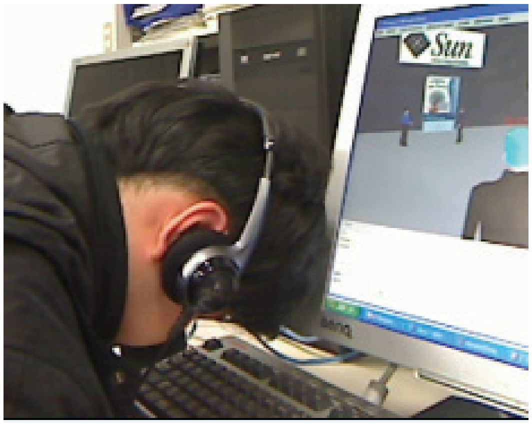

Figure 9. User B is sleepy and considered not hearing

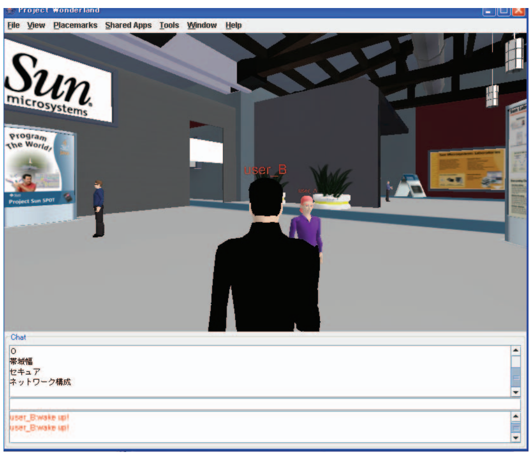

Figure 10. User B receives keyword/summary presentation after awaking by alert

[Test case 3]: Not hearing
Problem: User B is not hearing the speaker's explanation
System analysis:

- Webcam: head movement of user B (up/down head movement)
- Mouse sensor: Blood pulse frequency and skin temperature of user B.
- Voice interaction
- Acceleration meter

System action:
If no up/down head movement and chair movement, no voice signal from user B, and blood pulse frequency and skin temperature is decreasing
{Alert to wake up & Keyword/summary presentation}
Adaptive support:
User B can keep up with speaker's explanation.

Figure 11. Formal presentation of test case 3

V. Conclusions and Future Work

This paper proposed the adaptive keyword/summary presentation approach, which has been implemented and evaluated in the Wonderland Project for enriching the collaborative environment. The adaptive keyword/summary presentation features aim at identifying individual misunderstandings in remote collaborations and preventing potential misunderstandings. The identification of individual misunderstandings is based on various input parameters which help in assessing the personal situation of the users. If potential misunderstandings are detected, the users are supported with additional information about the current collaboration session by providing keywords and/or a summary of the current talk.

Our approach aims at reducing and preventing misunderstandings in remote collaborations and providing users with additional support in order to make remote collaborations a success. Future work will deal with evaluating the effectiveness and user-friendliness of our adaptive keyword/summary presentation features in collaboration sessions. Furthermore, we plan to develop and evaluate additional features with similar goals, namely to support users in overcoming misunderstandings in remote collaborations.

Acknowledgment

Authors acknowledge the support of Japanese Society for the Promotion of Science, iCORE, Xerox, Austrian Science Fund (FWF) under grant J2831-N13, and the research related gift funding by Mr. A. Markin, Research Center for Advanced Technologies of Tokyo Denki University, KAKENHI (21700244).

References

[1] F. Enembreck and J.-P. Barthes, Personal assistant to improve CSCW, *Proceedings of the International Conference on Computer Supported Cooperative Work in Design*, IEEE Press, 2002, pp. 329-335.

[2] K.C. Sia, J. Cho, and H.-K. Cho, Efficient Monitoring Algorithm for Fast News Alerts, *IEEE Transactions on Knowledge and Data Engineering,* 19 (7), 2007, pp. 950-961.

[3] RSS 2.0 Specification. Retrieved 24 March, 2009, from http://blogs.law.harvard.edu/tech/rss.

[4] RSScache. Retrieved 24 March, 2009, from http://www.rsscache.com, 2006.

[5] Y. Sakurai, S. Hashida, S. Tsuruta, and H. Ihara, Reliable Web-based CSCW Systems using Information Fusion of various multiple Biological Sensors, *Proceedings of the 4th International IEEE Conference on Signal Image Technology & Internet Based Systems (SITIS' 2008)*, Bali, Indonesia, 2008, pp.480-489.

[6] Y. Sakurai, Kinshuk, S. Graf, A. Zarypolla, K. Takada, and S. Tsuruta, Enriching web based computer supported collaborative learning systems by considering misunderstandings among learners during interactions. *Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT 2009)*, Riga, Latvia, 2009, in press.

[7] Sun Microsystems, Inc.: Project Wonderland: Toolkit for Building 3D Virtual Worlds. Retrieved 24 March, 2009, from https://lg3d-wonderland.dev.java.net/.

[8] A. Lee, T. Kawahara, and K. Shikano. Julius - an open source real-time large vocabulary recognition engine, *Proceedings of the European Conference on Speech Communication and Technology (EUROSPEECH)*, 2001, pp. 1691-1694.

[9] A.S. Tanenbaum and M. van Steen. "Distributed System Principle and Paradigms", Prentice Hall, 2002.