Discussion Mining: Annotation-Based Knowledge Discovery from Real World Activities

Katashi Nagao[1], Katsuhiko Kaji[2], Daisuke Yamamoto[3], and Hironori Tomobe[4]

[1] EcoTopia Science Institute, Nagoya University
nagao@nuie.nagoya-u.ac.jp
[2] Department of Media Science, Graduate School of Information Science,
Nagoya University
kaji@nagao.nuie.nagoya-u.ac.jp
[3] Department of Media Science, Graduate School of Information Science,
Nagoya University
yamamoto@nagao.nuie.nagoya-u.ac.jp
[4] 21st Century COE Program on Intelligent Media Integration,
Nagoya University
tomobe@nagao.nuie.nagoya-u.ac.jp

Abstract. We present *discussion mining* as a preliminary study of knowledge discovery from discussion content of offline meetings. Our system generates minutes for such meetings semi-automatically and links them with audio-visual data of discussion scenes. Then, not only retrieval of the discussion content, but also we are pursuing the method of searching for a similar discussion to an ongoing discussion from the past ones, and the method of generation of an answer to a certain question based on the accumulated discussion content. In terms of mailing lists and online discussion systems such as bulletin board systems, various studies have been done. However, what we think is greatly different from the previous works is that ours includes face-to-face offline meetings. We analyze meetings from diversified perspectives using audio and visual information. We also developed a tool for semantic annotation on discussion content. We consider this research not just data mining but a kind of real-world human activity mining.

1 Introduction

It will be very helpful that we can proceed to a discussion smoothly in a meeting or in a place for debating, and that we can retrieve and reuse the content such as minutes of these discussions. Generally speaking, however, unlike online discussions on bulletin boards on the Web, on offline ones at face-to-face meetings, making minutes reusable is difficult.

In this research, we propose a system that generates structured data on the discussion content semi-automatically and that displays them being visualized

in real time. In addition, the data are registered in an XML database and in a relational database so that retrieval and summarization will be available. To apply this function, we developed an experimental system for supporting discussions.

One of the purposes of our research is to acquire knowledge from read world human activities such as conversations and discussions. Also, we are researching methods to annotate human activity records like meeting minutes with some semantic structures and to reuse them for supporting ongoing activities.

We have been developing some technologies which can record the sounds and the visions in meetings and generate multimedia minutes semi-automatically. Our technologies also include search and summarization methods for finding and browsing relevant and useful information.

In this paper, we focus on two technologies. One creates discussion content not just meeting minutes but hypertext documents including participants' statements linked with video and audio. The discussion content also includes discussion graphs that visualize semantic/pragmatic structures of discussions. The other allows the users to annotate the discussion content with additional information about the structure of the discussion. The semantically-annotated discussion content is easily converted into more appropriate formats to be reused.

2 Creation of Discussion Content

There is a method mainly using video (image and sound) to make a discussion reusable. What we employed is a method that generates data on minutes primarily by text input through the forms in a Web document. The forms are subdivided and not only the content of the discussion but the information such as presenter, date, title and participants are inputted.

We especially focused on information about presentation slides in the trial system. In recent years people tend to use slides made by Microsoft Power-Point for meetings. We employed JACOB (A Java-COM Bridge) to handle COM (Component Object Model) with Java. Specified with Microsoft PowerPoint from the ones made in advance, the slide is converted to GIF format and is displayed on an input page of minutes. Simultaneously the information of characters in the slide is obtained and is added to the minutes. Then the minutes are structuralized by relating with the slide and a statement.

Furthermore the information of the statement is input after being categorized into three types, question, answer, and comment. We limited the types of statement to the three in order to avoid troubles when inputting its content in plain text. We assumed that the information of a discussion consisted of three kinds of statement, question, answer, and comment. So when a subject changes to another, we realized a system that enabled to show a new form for input.

Figure 1 shows a display screen of inputting minutes.

gIBIS [2] seems to represent groupware by structural approach. This can display the structure of a discussion graphically to turn into facilitating the grasp of the content of minutes and encouraging effective statements.
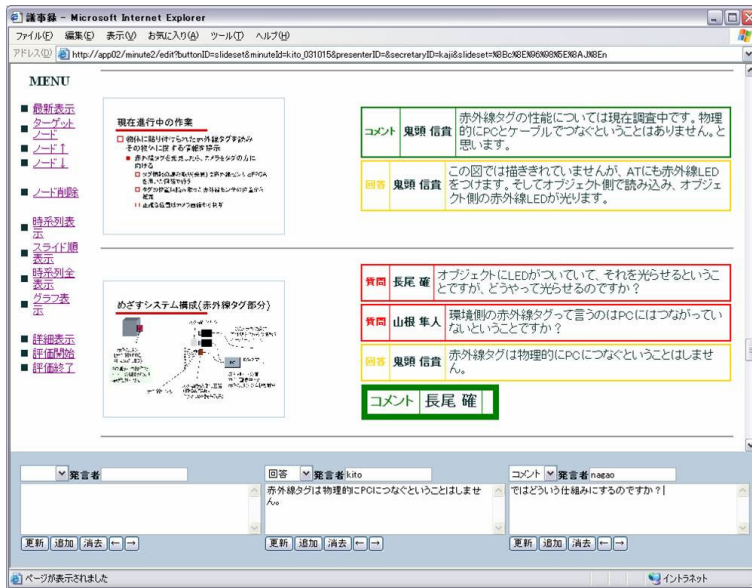
Fig. 1. Input of Minute

In this system, we visualized the structure of minutes by creating graphical display and edit mode of statements with the use of SVG (Scalable Vector Graphics) [6]. The graph was semi-automatically structuralized with the pertinent information and keyword on the statements and slide as shown in Figure 2. In result, it allows users to edit.

Minutes are mainly created from the inputted text and the presented slide. The format of minutes is XML (eXtensible Markup Language) and they are stored in an XML database and a relational database at a server machine with Java Servlet. On this study, we used Xindice [1] as the XML database and PostgreSQL as the relational database.

3 Reuse of Discussion Content

Some stylesheets on DOM (Document Object Model) object in Xindice allow real-time view of the minutes registered in the database through conventional Web browsers.

Additionally, the users can access and retrieve only required minutes from the search form on the Web as shown in Figure 3. As the retrieval field is subdivided into several fields for title, date, a presenter and so on, we can easily access, retrieve and view the minutes that we want. As far as the retrieval, high-speed access will be available thanks to PostgreSQL.
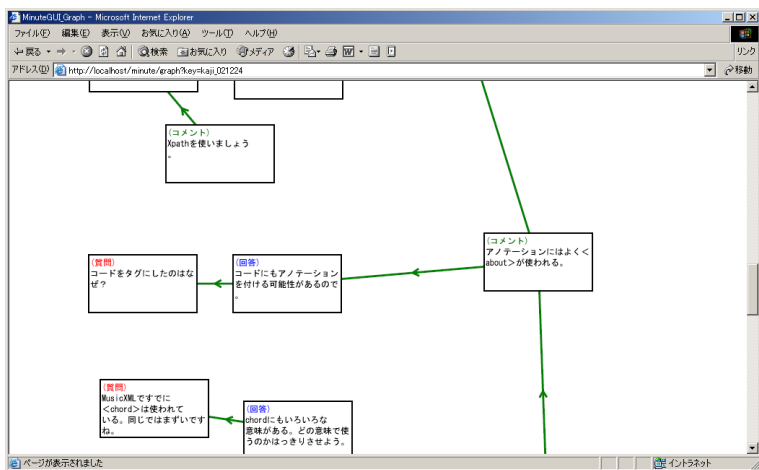
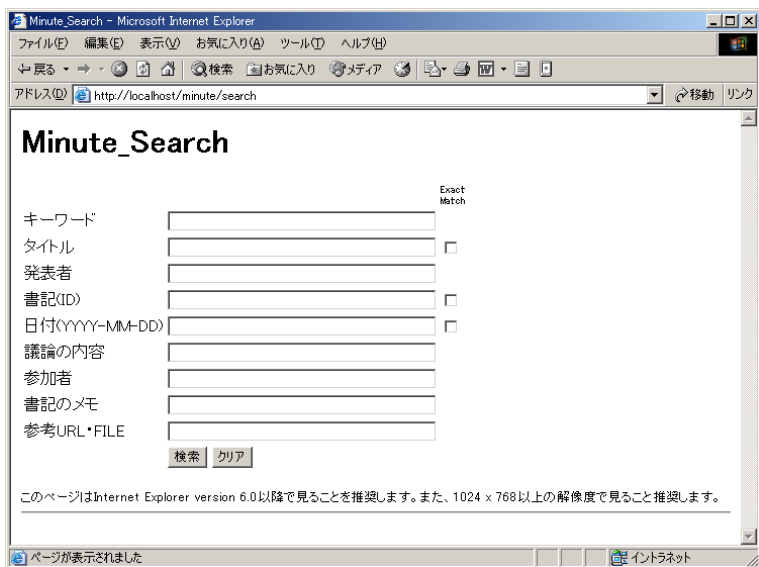**Fig. 2.** Graph View of Minute

Fig. 3. Minute Search

The users can reedit and add the minutes, and correct their inputting errors through the Web browsers. When the users need to see the main points of the minutes, they can also see the summarized versions of the minutes in the order of importance. We utilize the structure of the statements at the time of inputting, the keyword of the statements, and the earned points for grades added to manually that are described later.

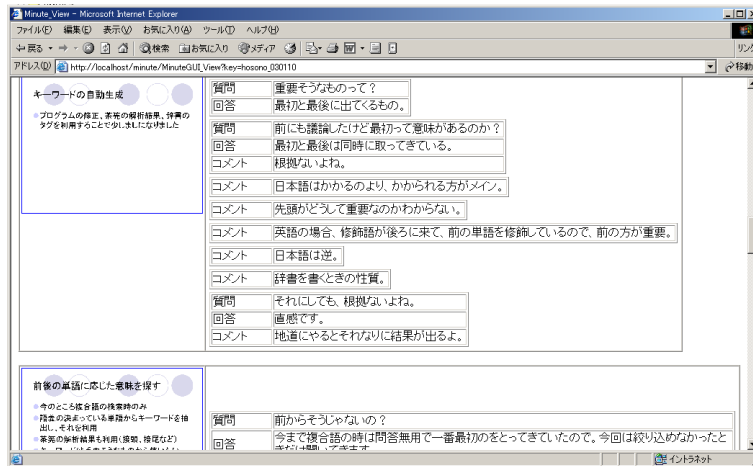Figure 4 shows an original version of a minute and its summarized version is shown in Figure 5.

Fig. 4. Summarization of Minute (before)

4 Semantic Annotation on Discussion Content

To utilize minutes more effectively, we add some semantic features to minute data such as linguistic structures, reference information to external content, and exact meanings of ambiguous words. We developed an authoring tool to create semantic annotation on minutes by linguistic analysis of text such as sentence parsing, anaphora resolution, and word sense disambiguation [5].

By using semantic information associated with mimutes, we can extract some important points in meetings and create summaries that include short history of discussions on the important topics.

Although annotation-based intelligent content processing is one of the major topics in some international research activities such as MPEG-7 [3] or Semantic Web [7], it has not obtained a definite result. We have already developed the system for interactive summarization and translation of Web content including multimedia data based on semi-automatically-created semantic annotation. Our annotation-based architecture is called "Semantic Transcoding" [5].
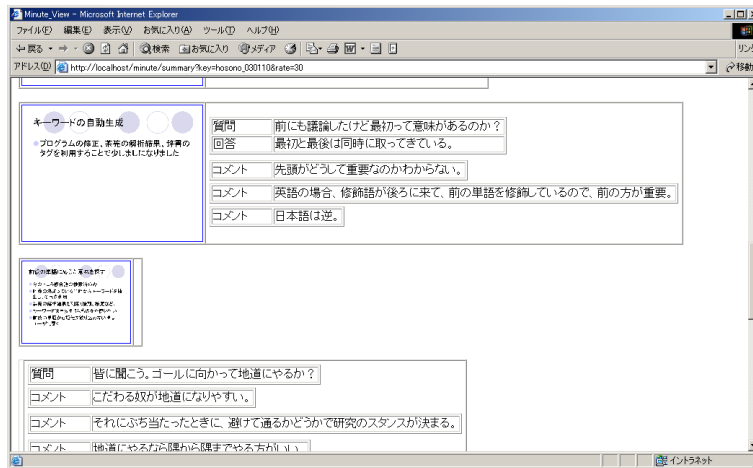
Fig. 5. Summarization of Minute (after)

The annotation here includes the detailed language structure of the sentences in the document and the semantic segments of multimedia content which are generated semi-automatically. Each piece of the annotation is correctly associated with some part of the content.

The annotation enables Web audiences use the content more easily by using the semantic relationships among some media. For example, searching for videos using a keyword and summarization of videos connected with the summaries of voice transcript texts became possible because of the annotation. However, generation of the annotation costs very much. We need some methods to make this task more cost effective without making the annotation generation more complicated and inaccurate. One of our solutions captures some explicit hints or suggestions which are given by human who can easily understand the meaning of content and infers the semantic structure underlining the content. Our new user tools and software will make these processes simple and easy.

5 Discussion Mining

Discussion mining is a new research domain that aims at digitization of discussion content and discovery of knowledge from the content. Our developed system records audio-visual scenes of discussions in face-to-face meetings and visualizes discussion content in the form of Web documents linked with videos captured by several cameras and microphones.

We have done some preliminary studies and evaluations of the prototype systems for discussion capture and visualization. We have also implemented a minute editing tool that is capable of semantic annotation.

The meeting participants support the automatic generation of the minutes by transmitting their user IDs and types of their statements at the meeting via InfraRed signals by using a special device called "the discussion tag". Each participant has three types of the discussion tag colored green, yellow, and red, respectively. The green one is indicated to all participants when the participant wants to make a comment. The red one is used for asking a question. The yellow one is for answering it. All tags are also used for evaluation of whole discussion. Our system detects IR signals from each participant's discussion tag and generates a graph structure consist of each statement (some keywords included in the utterance are manually inputted by a meeting secretariat) as a node and each relation between statements (question-its answer, referred comment-referrer comment, etc.) as an arc. The nodes are also linked with the corresponding scenes in the recorded video and audio.

Figure 6 shows a scene of the discussion using the system, and Figure 7 shows a screen shot of a structured minute generated semi-automatically. The minute is visualized as a graphical structure of the discussion, and allows the users to retrieve the video related to specific scenes. Figure 8 illustrates the whole system configuration, and Figure 9 illustrates the discussion tag system. How to use the discussion tag is very simple and easy. A potential speaker will hold up the discussion tag just before beginning of a talk at the meeting. The tag automatically transmits the data to the system.
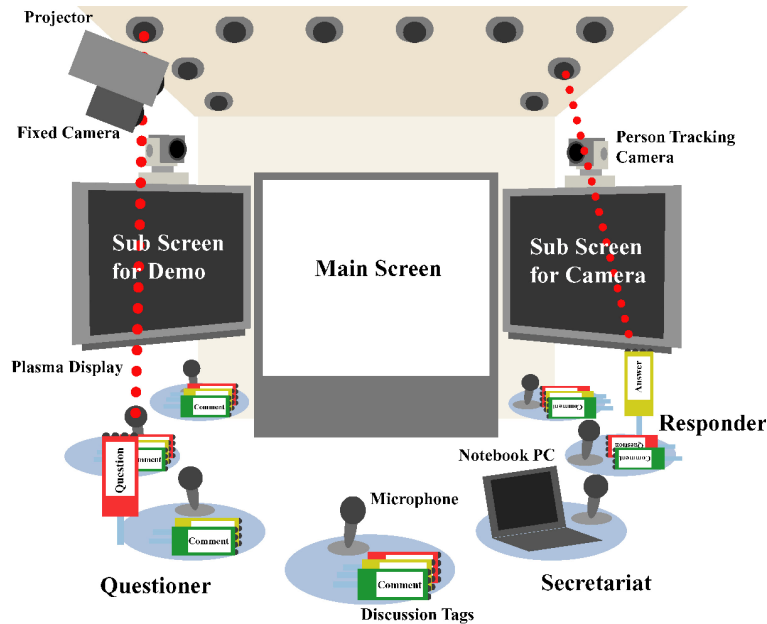

Fig. 6. Scene of Discussion

Fig. 7. Screen Shot of Structured Minute with Videos

Fig. 8. Configuration of Discussion Mining

As mentioned earlier, the content of the minutes is represented in an XML data format and stored in the XML database. The audio-visual content is accumulated in the multimedia database. These databases are connected with the network and the XML data include the pointers to the multimedia data.

The minutes in the XML format enable keyword-based retrieval and summarization. The summarization method is based on importance values calculated by node importance (more linked modes are more important) and word importance (more salient words are more important). We evaluate each discussion by using discussion tags (the green, yellow, and red tag indicate "good", "soso", and "not good", respectively). We also evaluate each statement at a meeting by using a Web-based discussion annotation tool. These evaluations are also considered to calculate the importance of the discussion.

Not only visualization, retrieval and summarization of single minute content, we have also developed a mechanism of visualization of multiple minutes. Our system performs grouping of the accumulated minutes and linking between related minutes. A group of the minutes includes similar minutes and their correspondences. The groups of the minutes are visualized as a "minute map" which shows distances among the minutes as shown in Figure 10. The map supports the users to observe the whole discussions deeply and to discover knowledge. In order to make the map more informative, we have to create the annotation that indicates semantic relationships among the minutes. By using such annotation,
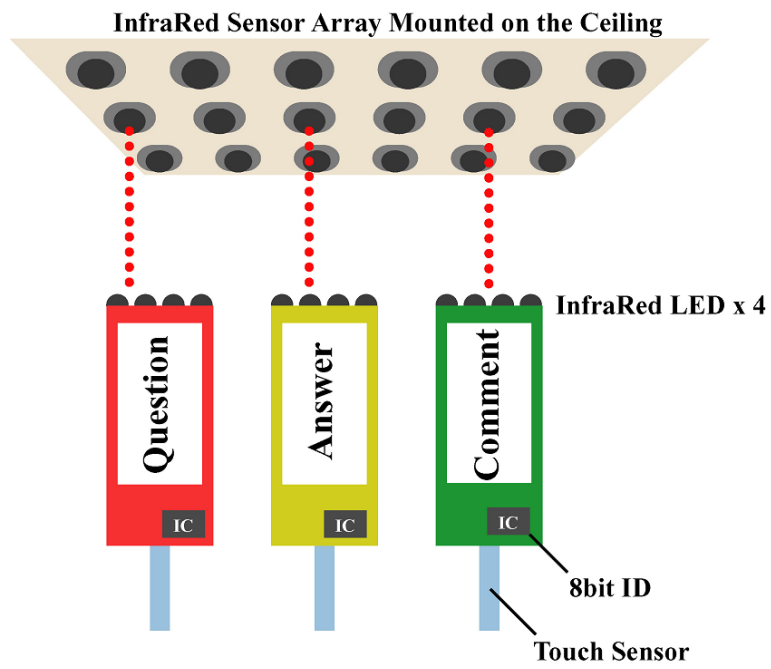
Fig. 9. Discussion Tag System

we can discover more detail about discussion structures across several minutes such as some background information of an argument and its derivations.
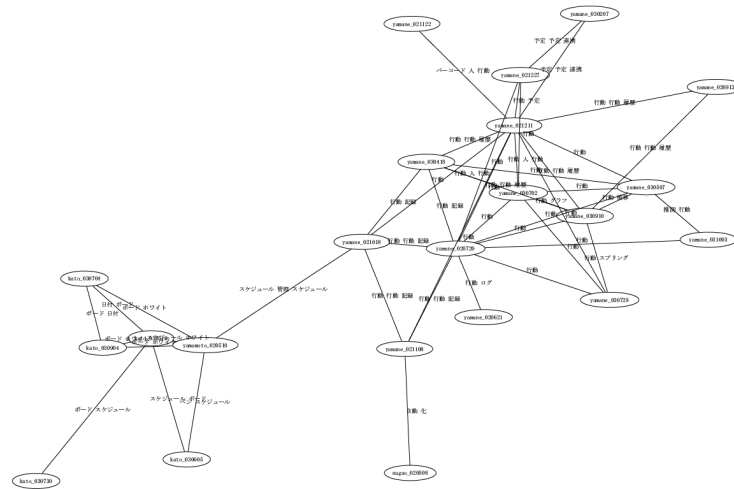

Fig. 10. Minute Map

From the minute map, we can also find the most important minute, the most significant statement and the most active speaker and questioner. Furthermore, we are analyzing the role of the sounds and the visions in meetings more deeply. We are also considering a new way of linking between multiple media related to human real-world activities such as conversations and discussions.

We should pursue some methods for not just capturing real world activities but also utilizing them for some meaningful tasks such as remembrance assistance and knowledge discovery.

6 Concluding Remarks

In this paper, we described the system of registering the minutes of meetings in database and reusing them. The following is the issues that we are planning to examine and to improve the system:

1. Annotation over Multiple Discussions
   Minutes generated by this system are not related to the past ones, for they only refer to a certain meeting. The system should be improved to examine relevance to the past minutes semi-automatically from a keyword and to summarize covering multiple minutes. It will help to create annotations across multiple minutes.

2. Use of Voice Transcripts

The use of voice will enable us to refer afterward to the information that cannot be contained in the minutes. Since the voice data is difficult to be searched as it is, using annotation can make it possible to largely reduce the problems like retrieval if speech recognition creates transcripts.

3. Use of Video Annotation

Since video data is available, the system will be able to utilize hard-to-get data from other information like mood of meetings and gestures and facial expressions of the participants. We have also developed a video annotation editor to annotate visual objects in video frames with some notes and attributes and to edit scene descriptions and voice transcripts [4].

In this research, we have shown one example of the knowledge discovery from the information in the real world, and the multiple media integration based on semantic annotation.

Our annotation system will be a strong tool to associate some elements in the media data with the other media elements. For example, the image and the motion of the speakers' reference expressions such as finger pointings will be connected with the particular linguistic expressions such as deixis in the voice and the text data. The relationships among the multiple media will be pointed out clearly by using the semantic annotation, and contribute to finding out the rule of them.

Furthermore, the integration of other knowledge sources with the semantic annotation on the minutes makes it possible to search for and add to the information about speakers who made good arguments.

Semantic annotation is supposed to cost much in generally. However, involving human intervention effectively into the annotation task when the original content or the material of content is created improves cost performance very much. Needless to say, improving the quality of automatic analysis of the content such as vision and speech recognitions is able to facilitate the efforts for creation of the semantic annotation.

We also need more user-friendly tools that reduce the human labor as much as possible. For example, in addition to the discussion tags, a mechanical button on the desk will be useful during meetings. Just pressing it by each participant lets the system know which comment is important in the discussion and which action is meaningful in the physical event. Such simple mechanism will work for automatic annotation to the audio-visual data.

Acknowledgments

References

1. The Apache XML Project. Apache Xindice. http://xml.apache.org/xindice/. 2001.
2. Conklin, J. and Begeman, M.L. gIBIS: A Hypertext Tool for Exploratory Policy Discussion. Proc. of CSCW '88. pp. 140–152, 1988.
3. MPEG. MPEG-7 Overview. http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm. 2002.
4. Nagao, K., Ohira, S., and Yoneoka, M. Annotation-Based Multimedia Annotation and Transcoding. in *Proceedings of the Nineteenth International Conference on Computational Linguistics (COLING 2002)*. pp. 702–708, 2002.
5. Nagao, K. *Digital Content Annotation and Transcoding*. Artech House Publishers. 2003.
6. W3C. Scalable Vector Graphics (SVG) 1.0 Specification. http://www.w3.org/TR/SVG/. 2001.
7. W3C. The Semantic Web Community Portal. http://www.semanticweb.org/. 2002.