Proceedings of the World Congress on Engineering and Computer Science 2018 Vol I
WCECS 2018, October 23-25, 2018, San Francisco, USA

Discussion Summarization Based on Hierarchical Structure Using Verbal and Non-Verbal Information

Hiroya Miura, Yoshinari Takegawa, Asuka Terai, and Keiji Hirata

***Abstract*—The aim of the study is to analyze the semantic relationships among utterances in a meeting at a deep level. To achieve the aim, we propose a novel tree structure, called discussion time-span tree, for representing a discussion structure based on the relative importance of utterances contained which are derived from the meaning of a meeting. A discussion time-span tree is a binary tree, each leaf of which corresponds to a utterance; it has the following characteristics: adjacent relevant utterances are hierarchically grouped, and a more important utterance becomes a head. Thus, a discussion time-span tree represents implicit meaning and intention contained in meeting records, and can be used for summarization. In the study, we propose a method and rules to generate a discussion time-span tree based on verbal and non-verbal information. In the evaluation on important utterance extraction using the proposed system, it was shown that the score of ROUGE-2 was higher than that of the existing document summarizing technique. Moreover, in the evaluation by subjects conducted by 10 subjects, we showed that the system can grasp conference contents efficiently, and demonstrated the usefulness of the system.**

***Index Terms*—automatic summarization system, discussion analysis, information retrieval**

I. INTRODUCTION

MINUTES, which record the content of the meetings, play a very significant role in the smooth advancement of work and research. However, it is difficult to grasp information that is not explicitly recorded in conventional minutes, as they often summarize superficial details [1]. Also, because recorded minutes contain how opinions were exchanged, what paths were taken until conclusions were reached, and many other pieces of information, they serve as content that yield new value through reuse. From these kinds of circumstances, minutes data sets with various annotations added by hand, such as AMI Meeting Corpus [2] and Discussion Mining [3] have been provided on the Web.

In general, the research on automatic summarization for these corpus uses existing document summarization technology. Concretely, there are many researches applying typical summarizing methods, such as LEAD-based Method [4] which extract opening lines from input document and LexRank [5] which is method based on the concept of PageRank [6]. Although these studies enabled to greatly reduce the amount of utterances to be read, there has been little consideration on the utterance structures or the correspondence among utterances, there are problems in the generation of syntactically unnatural sentences and meanings. Research is also taking place to construct arguments by analyzing discourse structures and meanings in order to rapidly discover and understand points made during meetings [7], [8], [9]. However, with these methods, although it is possible to acquire superficial group structures of discussion, but it is difficult to acquire hierarchical group structures. We consider that we can represent knowledge organized and understand semantic relationships, by hierarchical structure. On another front, many studies for discussion analysis suggest the importance not only linguistic information but also non-verbal information, such as utterance duration and turn-taking. Especially, research has been carried out from the 2000s to now regarding discussion analysis of the non-verbal communication of meeting participants' from the perspective of portraying meetings as multimodal interactions between many people [10], [11]. These research clarified effective feature quantities with high accuracy in discussion analysis. However, there has been little research focused on applying the results toward an automatic summarization system or the development of practical applications.

The aim of the study is to analyze the semantic relationships among utterances in a meeting at a deep level, targeting the discussion corpus recorded by Discussion Mining [3]. To achieve the aim, we propose a novel tree structure, called discussion time-span tree, for representing a discussion structure based on the relative importance of utterances containxed which are derived from the meaning of a meeting. A discussion time-span tree is a binary tree, each leaf of which corresponds to a utterance; it has the following characteristics: adjacent relevant utterances are hierarchically grouped, and a more important utterance becomes a head. Thus, a discussion time-span tree represents implicit meaning and intention contained in meeting records, and can be used for summarization. We consider that we can represent knowledge organized and understand semantic relationships, by hierarchical structure. The above idea of the tree structure has been inspired by music theory Generative Theory of Tonal Music (GTTM) [12], which is a theory that parses the chronological sequence of musical events. We can consider music theory as applying to analyzing a meeting record, paying attention to the hypothesis that both musical events in musical composition and utterances in discussion are generated some meaningful group (gestalt) along a time axis. In the study, we describe a method and rules to generate a discussion time-span tree based on verbal and non-verbal information, and evaluate its validity.

H. Miura is with RIKEN and Future University Hakodate, Tokyo 103-0027, Japan (e-mail: hiroya.miura@riken.jp). Y. Takagawa, A. Terai, and K. Hirata are with Future University Hakodate, Hakodate City, Hokkaido 041-8655, Japan.

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I
OUTLINE OF SAMPLE DISCUSSION

ID	Speaker	Utterance
U_1	O	I think avoiding a dangerous situation means looking at each other.
U_2	W	Recognizing each other is important.
U_3	O	They need to confirm the direction that a person moves in at least once.
U_4	W	Not being able to recognize each other is a problem.
U_5	N	If the robot cannot predict a person's next action, the robot cannot recognize and evade people.
U_6	W	If the robot does not recognize a person, it cannot predict an action.
U_7	N	What the robot should do is not only evade but also announce its presence.
U_8	W	Just as you say. It is necessary to be careful about such a thing.

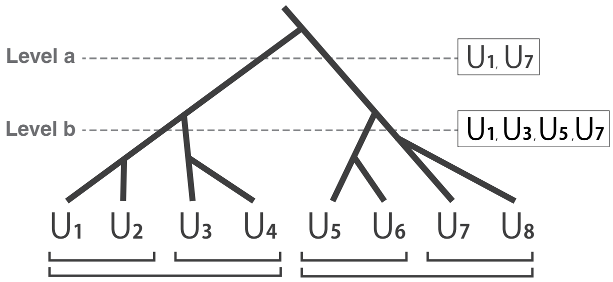

Fig. 1. Discussion Time-Span Tree

II. Discussion Time-Span Tree

A. Representation of Discussion Time-span Tree

In this study, we adopt a hierarchical, ordered structure as a method to represent hidden meaning and/or intention in meeting records. Data of this nature are time-series data. People comprehending the discussion structure from the relationships between utterances, e.g., whether one is for or against the other, or a question and an answer, which yield a gestalt in understanding the discussion. For adjacent utterances, one is more important than the other. Hence, the semantic structure of a discussion is modeled using a tree structure that represents the importance of the utterances hierarchically; we refer to it as a discussion time-span tree (Figure 1). By using this expression method, we can understand the discussion structure and the general flow of this discussion step by step based on the relationship between the utterances.

The discussion time-span tree is a binary tree that has a branch for each utterance, and it has the following characteristics: (i) it regards adjoining relevant utterances as one group; and (ii) it represents the hierarchical importance of the utterances. People understand meeting records and discussion structure based on the relation of each utterance to other utterances, e.g., whether it is for or against, and question or answer, which yield a gestalt in understand discussion. Table I shows an example of a thread that is composed of eight utterances. The discussion time-span tree is constructed on this thread. Regarding (i), U_1 and U_2 generate a subgroup at the lowest level, and the additional hierarchical structure is generated from the bottom up. Regarding (ii), the tree represents the hierarchical importance of utterances through the relationship of each branch in the generated tree structure. In Figure 1, the utterances are ranked in order of importance as U_7, U_1, U_3, ... down to U_2.

TABLE II
LIST OF GROUPING PREFERENCE RULES (GPR)

Item	Parameter
GPR1a	interval between utterance
GPR1b	order of speakers
GPR1c	change in the number of words in utterance text
GPR1d	utterance duration
GPR2a	presenter's utterance
GPR2b	first occurrence of important words

TABLE III
LIST OF SIGNIFICANCE PREFERENCE RULES (SPR)

Item	Parameter
SPR1a	utterance duration
SPR1b	number of words in utterance text
SPR1c	amount of approval
SPR2a	start-up and last utterance
SPR2b	social status
SPR2c	presenter's utterance
SPR3a	first occurrence of important words
SPR3b	utterance contains important words

B. Generation Method of Discussion Time-span Tree

A discussion time-span tree is generated in the bottom-up manner from the information according to two kinds of rules: (1) rules that acquire grouping structures in the discussion (Grouping Preference Rule: GPR); and (2) rules that select a significant utterance that represents the duration of a certain entire group (Significance Preference Rule: SPR). In this research, we propose GPRs and SPRs with reference to [10] and [11] which clarified effective feature quantities with high accuracy in discussion analysis. Tables II and III presents the list of GPRs and SPRs, respectively. These rules can judge the similarity of topics based on temporal proximity, speaking order, and text information. In addition, we assume that the importance of an utterance can be decided by the number of times it is uttered, the duration, and the frequency of occurrence of important words. Regarding (i), for example, GPR 1a could be: "Consider a sequence of four utterances, n1-n4; the transition n2-n3 may be considered a group boundary if the interval between utterance is changed." Regarding (ii), for example, SPR 3b could be: "Consider a sequence of four utterances, n1-n4; the n1 may be considered more significant if the utterance which contains important words."

Since GPRs mix rules on global structure and local structure, it is difficult to properly execute both rules. In order to deal with this problem, we have designed an algorithm

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

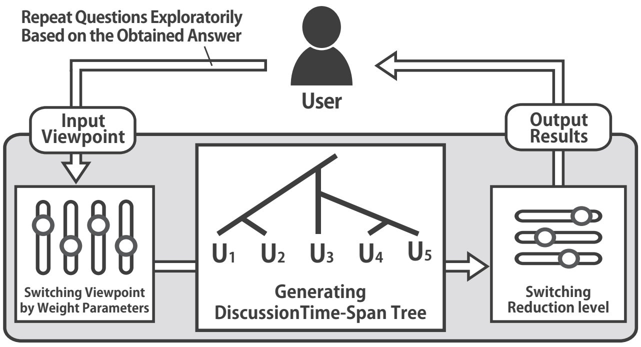

Fig. 2. System Diagram

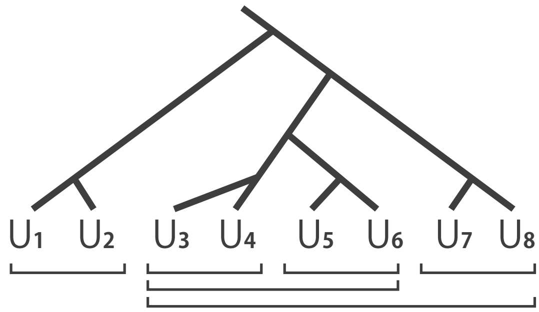

Fig. 3. Viewpoint Switching of Discussion Time-Span Tree

that appropriately combines local and global processing. The procedure for generating a discussion time-span tree is shown below. First, for the hierarchical group structures, it is necessary to acquire the top-down using local boundaries determined through bottom-up processing. Therefore, we compute the strength of the higher order of the boundary by the boundary value between utterances according to the applied GPRs. We consequently split the group into two chunks from the strongest boundary, and repeat the process if the group contains an internal local boundary. Thus, a local / global hierarchy can be obtained. The discussion time-span tree is generated bottom-up from these local / global hierarchies and the degree of the importance of each utterance according to the applied SPRs to the whole group.

C. Setting of Weight Parameters

We designed the parameters that manipulating the rules weights, called weight parameter, for generating discussion time-span tree. These parameters are weighted by considering the number of applications of each rule, normalizing in the range of values (0.0 to 1.0), and adjusting the weight value (the default value is 0.5). We designed the user interface of this parameter as slider GUI. The final boundary value is obtained by the multiplication of the boundary value given by applying rules and the weight value set by the parameter. The system users can improve analysis accuracy and generate a tree structure corresponding to different viewpoints, by operating this weight parameter freely. The operation of this weight parameter is an important task in giving new viewpoints for information retrieval. However, the users need to operate various weight parameters with trial-and-error, since each parameter has a mutual relationship.

III. System Overview

In this section, we describe a system for structuring discussion and generating summary interactively, by using the mechanism based on discussion time-span tree. Figure 2 shows the system diagram. Exploratory data analysis of discussions requires a trial-and-error process because the information required by the user is initially undefined, and only becomes clearer during the data analysis process. The problem is solved by repeating this exploratory operation. The user can continue exploring by structuring discussion and generating summary, based on the answer they obtain. A series of these operations allows the user to satisfy their information needs. The main functions of the system are the following two items.

Summarization by Reduction

The reduction of a discussion time-span tree assigns structural importance to each utterance in a hierarchical manner. The reduction is identified with the subsumption relation, which is known as the most fundamental relation in knowledge representation. Figure 1 shows the reduction concept. The use of reduction enables us to produce a group of utterances that summarize the original meeting records. For instance, cutting at the upper level (level a) generates a summary composed of two utterances, U_1 and U_7. Cutting at the lower level (level b) generates a summary composed of four utterances, U_1, U_3, U_5, and U_7. This mechanism is effective for understanding incremental discussion information. In the present function, it is possible to switch the reduction level by selecting an arbitrary number of utterances.

Viewpoint Switching by Parameter Adjustment

The adjustable parameters alter the weighted value of the rules for extracting intended important utterances and structuring the information for different purposes and from different viewpoints. The discussion time-span tree in Figure 1 is generated by giving importance to utterances such as question (U_1 and U_5) and opinion (U_3 and U_7).

On the other hand, the discussion time-span tree in Figure 3 is generated by giving high weight values to GPR1b, GPR1b, SPR2a, and SPR2c. Regarding SPR, for instance, giving priority to SPR2a, the first and last utterances are of high importance, and priority for SPR2c means that giving importance to presenter's utterances (W is the presenter in Table I). The summary that is generated ranks the utterances of raising a problem and introducing a new topic (U_1, U_3 and U_8) more highly. In this way, several different summaries can be generated.

IV. Experiment Results and Analysis

The purpose of proposed system is to grasp the summary efficiently by using the mechanism of the discussion time-span tree. Therefore, we conducted evaluation experiments on the following three points; (1) Evaluation on extracting important utterances from meeting records: Evaluation on whether important utterances can be taken out from the meeting records accurately. (2) Usefulness evaluation of the system by subject experiment: Evaluation on which of the following two conditions can better grasp meeting contents more efficiently; (x): in the case of using the proposed system, (y): in the case of viewing a web browser which

TABLE IV
SETTING VALUES OF WEIGHT PARAMETERS OF GROUPING PREFERENCE RULES (GPR)

1a	1b	1c	1d	2a	2b
0.7	0.5	0.3	0.5	0.4	0.2

TABLE V
SETTING VALUES OF WEIGHT PARAMETERS OF SIGNIFICANCE PREFERENCE RULES (SPR)

1a	1b	1c	2a	2b	2c	3a	3b
0.3	0.3	0.6	0.8	0.5	0.6	0.4	0.3

displayed only utterance text information. (3) Analysis on the method of reproducing the viewpoint for information retrieval: Analyze the usage tendency of the weight parameter operation in the subject experiment described in the previous section.

A. Discussion Corpus

In this experiment, we used the meeting record[1] which is transcribed text data of a utterance recorded with a technique called Discussion Mining and including its meta information as corpus. The objective of this study was to record real world data as text, video, meta data and annotation, and extract reusable data for meeting activities [3]. The system developed in this study supports the kind of meeting in which a speaker delivers a presentation with slides. In addition, the corpus divides into the continued question and answer segments with participants.

Now, let us call an utterance which presents a new topic as a start-up utterance, and an utterance which concerns the same topic as the previous one as a follow-up utterance. One start-up utterance and at least one follow-up one form the unity of each topic. We call it a *discussion segment*. The root of the discussion segment is the start-up utterance, and the rest is composed of follow-up utterances. A discussion time-span tree is generated for each discussion segment.

In the implementation of important words in GPR2b, SPR3a, and SPR3b, we obtained the part-of-speech information from the text data of utterances, by the morphological analysis. We used MeCab[2] for morphological analysis, and extracted nouns and adjectives included in each discussion segment. Finally, we introduced the top 10% of the words obtained by applying the TF-IDF method as important words.

B. Evaluation on Extracting Important Statements from Meeting Records

In this experiment, we conducted an evaluation on extracting important utterances from meeting records. We compared with methods commonly used in the relevant field, in order to verify whether extracting of important utterances by discussion time span trees is effective. Specifically, we used the following two methods in this experiment.

LEAD-based Method

[1]Nagao Laboratory of Nagoya University: Discussion Mining Project, Source ⟨http://dm.nagao.nuie.nagoya-u.ac.jp/⟩.
[2]MeCab ⟨http://taku910.github.io/mecab/⟩.

TABLE VI
EVALUATION OF IMPORTANT UTTERANCE EXTRACTION BY ROUGE-2

Method	ROUGE-2
LEAD-based method	0.318
LexRank	0.362
Proposed method	0.401

Text that appears near the beginning of a document tends to contain important information, and so this method is a means of selectively extracting these particular texts. In this experiment we extract opening lines from discussion segments until a specified summarization criterion is met.

LexRank

This is a classic summarization technique based on the concept of PageRank that uses a graphical representation of Erkan et al.'s sentence similarity measure. This method involves: (a) calculating the similarity between statements in discussion segments using TF-IDF and creating a similarity graph with the statements as nodes and the relationships between them as edges, (b) creating an adjacency matrix where when the degree of similarity is above the threshold, this is denoted by 1, and is denoted by 0 otherwise, and (c) calculating the principal eigenvector of the adjacency matrix of the aforementioned graph, and then proceeding to sequentially extract statements in order of decreasing node importance until a specified summarization criterion is met.

We conducted experiments on 25 discussion data recorded by Discussion Mining (Total discussion duration: 48 hour 36 minutes, number of discussion segments: 339 segments, number of utterances: 1661 utterances). We used ROUGE [13], which is most widely used method in automatic summarization evaluation, as an evaluation index. The measure is computed by counting the number of overlapping words or N-grams between the system-generated summary to be evaluated and the reference summaries. We adopted ROUGE-2 which is the most widely used. The reference summaries was independently annotated by two authors of the paper who are familiar with the generating method, and the inter-annotator agreement between the two was 68%. We set that the summarization rates are about 50%.

The proposed technique uses discussion time-span tree corresponding to the output that serves as the baseline determining the value of each weight parameter. Each weight parameter needs to be assigned the optimum value in order to reproduce reference summaries, and so a preliminary experiment was carried out in order to assign such values to them. This preliminary experiment used 20 meeting and conference minutes, excluding the minutes used for the main experiment, as training data. One of the experimenters spent 2$\sim$5 minutes on each data set adjusting the values while at the same time referring to the meeting records and reference summaries, such that the system output resembled the summaries. The weight parameters were set to be equal to the median of the values assigned to them during the preliminary experiment (Table IV and V).

Table VI shows the experimental result. Based on the results above, the proposed method is seen to produce results closer to reference summaries than the other standard sum-

TABLE VII
RESULTS OF THE COMPREHENDING TEST ON EACH SUBJECT

Subject	Number of Correct Answers Condition (x)	Condition (y)
Subject A	4	2
Subject B	7	3
Subject C	6	5
Subject D	5	2
Subject E	6	4
Subject F	6	2
Subject G	5	3
Subject H	7	3
Subject I	6	5
Subject J	5	2

TABLE VIII
SETTING VALUE OF WEIGHT PARAMETER AT THE END OF TASK FOR EACH SUBJECT

Subject	A	B	C	D	E	F	G	H	I	J
GPR1a	1.0	1.0	1.0	0.8	0.5	1.0	1.0	0.8	0.0	0.6
GPR1b	1.0	1.0	0.5	1.0	0.8	0.0	1.0	0.0	1.0	0.4
GPR1c	1.0	1.0	1.0	0.5	0.5	0.0	0.5	1.0	1.0	0.0
GPR1d	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.4	0.2	0.7
GPR2a	1.0	0.0	0.5	0.8	0.5	1.0	1.0	0.3	0.0	0.3
GPR2b	0.5	0.0	0.5	0.5	0.8	1.0	0.5	0.5	0.0	0.0
SPR1a	1.0	1.0	1.0	0.8	1.0	1.0	1.0	0.0	0.2	0.0
SPR1b	1.0	1.0	1.0	0.8	0.8	1.0	0.5	0.0	0.0	0.4
SPR1c	0.5	1.0	1.0	1.0	0.8	1.0	1.0	0.6	0.0	0.0
SPR2a	1.0	0.0	0.2	0.5	0.5	0.0	1.0	0.0	0.0	0.6
SPR2b	1.0	0.0	1.0	0.2	0.5	1.0	0.5	1.0	0.5	0.0
SPR2c	0.5	1.0	1.0	0.4	0.0	1.0	0.5	1.0	0.0	0.7
SPR3a	1.0	1.0	1.0	0.8	0.5	0.0	1.0	1.0	0.0	0.0
SPR3b	1.0	1.0	0.5	0.8	0.5	1.0	1.0	0.0	1.0	0.8

marization technique that makes use of comparative methods. This is thought to be because the comparative method does not take into account the structure of a discussion and forms a summary solely based on verbal information, resulting in a lower score than the proposed method. The results could be improved by taking into account not only verbal information but also non-verbal one. According to the above results, we claim that summarization method using discussion time-span tree is highly effective. On the other hand, there is a high possibility that fluctuation will occur in the analysis, because even a slight change of manipulating the weight parameter tends to change the result greatly. For this reason, it may be concluded that in order to increase the accuracy of summaries there is a need to propose a method for assigning and controlling values that takes into account the particular characteristics of each meeting records.

C. Usefulness Evaluation of Proposed System by Subjects Experiment

We conducted a test to verify the users' comprehension degree, in order to evaluate usefulness of the system. In this test, the subjects answered questions about the contents of the discussion under the following two conditions; (x): in the case of using the proposed system, (y): in the case of viewing a web browser which displayed only utterance text information.

The answers to questions prepared beforehand in these experiments was selecting one of four choices, and the subjects can answer correctly if they read all important utterances. The test consisted of the eight questions, and response time to questions was less than 10 minutes. The subjects were 10 students in their twenties, who majoring in computer science. Furthermore, in condition (x), a five-minute practice session was held prior to the experiment for the subjects to get a feel for how to use the system to perform functions such as adjusting weight parameters and so on. Each subject also individually evaluated different, separate meeting records under their respective conditions. The initial value of the weight parameter is as shown in Table IV and V.

Out of the 25 meeting records available online, 2 were randomly selected [3] and were combined with different conditions with each subject, in this experiment. Table VII shows the results list of the comprehension degree test. As a result of t-test with a significance level of 5%, the null hypothesis of no difference in number of correct answers between (x) and (y), was rejected. From the above, we show that the condition using the proposed system is significantly effective.

[3] ⟨http://dm.nagao.nuie.nagoya-u.ac.jp/guest/view?key=kawanishi_111021⟩. and ⟨http://dm.nagao.nuie.nagoya-u.ac.jp/guest/view?key=nagao_111026⟩.

D. Analysis on Method of Reproducing Viewpoints for Information Retrieval

In this section, we describe the analysis on the usage trend of the weight parameter operation in the subject experiment described in the previous section. The operation of this weight parameter is an important task in giving new viewpoints for information retrieval. However, the users need to operate various weight parameters with trial-and-error, since each parameter has a mutual relationship. Hence, we needed to verify how the subjects manipulated the parameters and relationship between the parameters, from the results of the experiments. Table VIII is the setting value of the final weight parameter at the end of the task in the subjects experiment. All subjects manipulated this weight parameter on average about four times. We regard that these setting values are noticeable as a result of trial-and-error in parameter manipulation of each subject.

We carried out a principal component analysis of the values of these weight parameters. The principal component analysis is a multivariate analysis to reduce the number of dimensions in data by obtaining the eigenvalues of the data matrix, allowing easy knowledge extraction from data distributions. The input data consists of the final values assigned to the weight parameters (0.0$\sim$1.0) by the subjects at the end of the task. We carried out a principal component analysis of the data set consisting of the values assigned to the weight parameters by all 10 subjects, and computed the contribution and cumulative contribution ratios of each component, the factor loadings for each variable with respect to the components, and the score plots for each principal component by each subject.

According to the analysis, the contribution ratios of each principal component were 28.1% for Component 1, 27.2% for Component 2, and 17.6% for Component 3, and the cumulative contribution ratio was 72.8%. This indicates that out of the 14 total parameters, over 70.0% of the information is contained within the first 3. Table IX shows the factor loadings of principal components 1$\sim$3 with respect to the values of each weight parameter. Moreover, the underlined

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE IX
FACTOR LOADINGS OF COMPONENTS 1∼3 TO WEIGHT PARAMETER VALUE

-	Component 1	Component 2	Component 3
GPR1a	-0.03	0.40	0.33
GPR1b	-0.46	-0.19	0.07
GPR1c	-0.40	0.15	0.09
GPR1d	-0.02	0.43	0.17
GPR2a	0.24	-0.17	0.45
GPR2b	0.46	-0.20	0.06
SPR1a	0.08	0.10	0.14
SPR1b	0.09	0.35	-0.03
SPR1c	0.07	0.15	-0.40
SPR2a	-0.17	-0.32	0.44
SPR2b	0.31	0.06	0.37
SPR2c	0.07	0.50	0.01
SPR3a	-0.46	0.04	0.15
SPR3b	-0.03	0.16	0.32

parts indicate moderately relevant parameters (r=0.31∼0.50). It can be seen from Table IX that GPR1a contributes to principal components 2 and 3, SPR2a to 2 and 3, and SPR2b to 1 and 3, with each contributing to more than one principal component. This indicates the effectiveness of these parameters. Finally, the principal components and their relevant weight parameters along with their defining characteristics may be divided as follows: Component 1 with verbal information (such as GPR1c, 2b, SPR3a), Component 2 with temporal information (such as GPR1a, 1d, SPR2a), and Component 3 with social signals information such as those indicating agreement or the strength of social influence (such as GPR1c, SPR2b) (Table IX). This suggests that the subjects had a tendency to search for information based on these three perspectives.

V. Conclusion

In this paper, we developed a summarization system of meeting records based on hierarchical structure. To realize the system, we presented a data model for representing hierarchical discussion structure by the discussion time-span tree, based on verbal and non-verbal information. By using this model, we aimed to be able to build an automatic summary of a meeting record, corresponding to different viewpoints. From the results of evaluation on extracting important utterances, it was shown that the score of ROUGE-2 was higher than that of the existing document summarizing technique. The results could be improved by taking into account not only verbal information but also non-verbal one. Moreover, from the usefulness evaluation of the system by subject experiment, it was found that the system users improved their comprehension degree significantly, compared to when not using the system. Generally, the system users were able to reach the same comprehending regarding the flow of the meeting and the important arguments. From these results, it was demonstrated that the proposed system could efficiently grasp discussion contents, and corroborated the utility of the discussion time-span tree which represents hierarchical discussion structure. Finally, from the analysis on the usage tendency of the weight parameter operation in the subject experiment, it was found that the subjects had a tendency to search for information based on following three perspectives; verbal information, temporal information, and social signals information.

Future work will also consider the relevance of a general formulation. Since, the authors propose a system for algebraic operations (such as join and meet) on a time-span tree, which is an analysis result of music [14], we aim to render meeting records reusable for a variety of purposes, by applying the operations, such as join and meet, to the discussion time-span tree. In addition, we will implement the function of assigning to the rule parameters the default values for a specific retrieval mode. Moreover, it is necessary for us to carry out large-scale experiments that increase the number of subjects and the task variationally, and classify their usage trends.

Acknowledgment

This research received a grant from JSPS Grant-in-Aid for Scientific Research 16H01744.

References

[1] Arpit Sood, "Towards Summarization of Written Text Conversations. Doctoral Dissertation", International Institute of Information Technology, 2013.
[2] Jean Carletta, Simone Ashby, Sebastien Bourban, et al., "The AMI meeting corpus: A pre-announcement", In Proceedings of MLMI, the Machine Learning in Medical Imaging pp.28-39, 2005.
[3] Katashi N., Katsuhiko K., Daisuke Y., and Hironori T., "Discussion mining: Annotation-based knowledge discovery from real world activities", In Proceedings of PCM, the Fifth Pacific Rim Conference on Multimedia, Springer-Verlag Berlin Heidelberg, pp.522-531, 2004.
[4] Edmundson, H., "New methods in automatic abstracting", Journal of the ACM, Vol.16, No.2, pp.264-285, 1969.
[5] Erkan, G. and Radev., D., "Graph-based lexical centrality as salience in text summarization", Journal of Artificial Intelligence Research, Vol.22, p.457-479, 2004.
[6] Page, L., Brin, S., Motwani, R., and Winograd., T., "The PageRank Citation Ranking: Bringing Order to the Web", In Proceedings of WWW, the 7th International World Wide Web Conference, 1998.
[7] Annie L., Aravind J., and Ani N., "Discourse indicators for content selection in summarization", In Proceedings of SIGDIAL, the Annual Meeting of the Special Interest Group on Discourse and Dialogue, pp.147-156, 2010.
[8] Hirao, T., Yoshida, Y., Nishino, M., Yasuda, N., and Nagata, M., "Single-document summarization as a tree knapsack problem", In Proceedings of EMNLP, the Conference on Empirical Methods on Natural Language Processing, pp.1515-1520, 2013.
[9] Mohamed E. and Ahmed E., "Text to multi-level MindMaps, Multimedia Tools and Applications", Springer, pp.1-28, 2015.
[10] Junko Ichino, "Discriminating Divergent/Convergent Phases of Meeting Using Non-Verbal Speech Patterns", In Proceedings of ECSCW, European Conference on Computer-Supported Cooperative Work 2011, pp.153-172, 2011.
[11] Okada, S., Ohtake, Y., Nakano, Y., Huang, H., Hayashi, Y., Takase, Y., and Nitta, K., "Estimating communication skills using dialogue acts and nonverbal features in multiple discussion datasets", In Proceedings of ICMI, the International Conference on Multimodal Interaction, pp. 169-176, 2016.
[12] Lerdahl, F., and Jackendoff. R., "A Generative Theory of Tonal Music", The MIT Press, 1983.
[13] Lin, C. and Hovy, E., "Automatic Evaluation of Summaries Using N-gram Co-occurrence Statistics", In Proceedings of NAACL, the 4th Meeting of the North American Chapter of the Association for Computational Linguistics and Human Language Technology, pp.150-157, 2003.
[14] Tojo, S. and Hirata, K., "Structural similarity based on time-span tree", In Proceedings of CMMR, the International Symposium on Computer Music Modeling and Retrieval, 2012.