Regular Paper

Automating Viewers' Side Annotations on TV Drama from Internet Bulletin Boards

Hiroshi Uehara[†] and Kenichi Yoshida[†]

This paper proposes a method for creating the viewers' side annotations that reflect viewers' attentions on TV dramas. Internet bulletin boards are filled with large amount of viewers' dialogues concerning TV programs. Our approach is to extract the viewers' attention embedded in these dialogues and express them in the form of the graphical structures called attention graphs. The attention graphs act as viewers' side annotations. Viewers' side annotations assist in providing viewers with hints to locate their favorite scenes in full-length TV programs. In general, Internet bulletin boards are described without any particular order and are expressed in a poor grammatical manner. Our approach is to statistically recognize the notable frequencies of the words which represent viewers' attentions out from such unstructured documents. Attention graphs are the outcome of this approach. Attention graphs are evaluated by three types of tests. The test results demonstrate that attention graphs sufficiently act as viewers' side annotations, in terms of pointing out which scene the viewers pay attention to and clarifying how deeply viewers are impressed by the scenes.

1. Introduction

The rapid increase in the capacity of local storage devices such as hard disk recorders has enabled to record a whole month of TV programs. On the other hand, spare time to watch TV is not expected to increase in daily life. As a result, the gap between the large volume of stored TV programs and viewers' spare time for watching TV is being enlarged. One possible solution to bridge this gap is to make use of annotations which act as indexes to each segmentised period within a full-length program, thereby one can pick out only the highlights. However, constructing annotations on each program manually is a time consuming and expensive process. Thus a way to automate annotations is required.

Several related works presented ways to automate annotations. The approach presented in these works is to make use of the documents derived from the production of TV programs, such as scripts or scenarios. In general, these documents are described in a clear step wise manner in accordance with the progress in corresponding TV programs. This characteristic can be effectively used to find out the point where the annotations should be attached in the TV programs. Thus, the use of these documents leads to successful automated annotations. Nevertheless, the annotations are not necessarily ef-

fective in the sense that they might not provide viewers with hints on the scenes occupying viewers' interests, because the annotations are based on documents that reflect the producers' intentions, but not viewers' attention.

Internet bulletin boards and weblogs are filled with dialogues concerning TV programs. If we could apply the same approach above to these documents, the created annotations might reflect viewers' attentions because they are completely on the viewers' point of view. However, their ambiguous description styles prevent previous approaches from automating annotations, because these approaches work only when step wise logical documents exist.

In this paper, we try to find statistical regularities out of a large amount of unstructured dialogues in the Internet bulletin board to create "viewers' side annotations". What we call "attention graph" is a representation of viewers' side annotations that reflect viewers' attentions. The idea of creating attention graphs have been introduced in Ref. 1) with a simple qualitative test. In this paper, we apply the idea to drama programs and evaluate the performance of attention graphs quantitatively and graphically. Thus, we show how effectively the attention graphs act as the viewers' side annotations.

Section 2 of this paper first surveys related work and points out its limitations, and clarify the motivation of this research. Section 3 explains the basic ideas, and Section 4 reports

† University of Tsukuba

on the implemented system. Section 5 explains the experimental results. Finally, Section 6 concludes our findings.

2. Related Work

In prior research, several different approaches have been made to automate annotations on TV programs. One approach is to make use of text documents as introduced in the previous section. One study following this approach has applied the scorecard as the text document to annotate baseball games [2]. Typically, baseball games consist of repeated events, defined as the batting performance of each player. For example in one event, a player makes a hit to gain a base. In another event a player gets struck out. A scorecard has a clear repeating patterns of such events as observed in the game. This research has made use of this regularities and successfully create the annotations to baseball games on the air. A similar study can be found in case of TV cooking programs [3]. In this case, the research focused on the stepwise pattern found in the expository text books for cooking programs and has successfully created annotations called "flow graphs" that synchronizes with its cooking procedure. Another study is concerning news programs. It has attempted to annotate news broadcasting programs based on captions [4],[5]. Drama programs have also been a theme for the same type of research. Acting scripts have been applied in this case [6]. Generally, all these type of approaches depend on the derivational documents described in a clear stepwise manner.

A different approach is to make use of acoustic characteristics or iconic characteristics found in TV programs themselves. One example in the research made use of iconic characteristics in terms of the changing patterns of shots in programs. The research claimed that highlights in the programs are generally accompanied by more frequent changes of shots than the other scenes. Based on these findings, the study proposed a way to extract highlighted scenes out of the TV programs [7]. The efficiency of these kinds of approaches is limited, because each scene in a program is not necessarily accompanied by iconic characteristics or acoustic characteristics.

We try to automate annotations by making use of Internet bulletin boards. Our approach is similar to those described above in terms of making use of text documents. Nevertheless, it

can be said to be an unprecedented approach in that we intend to automate viewers' side annotations which reflect viewers' attentions by making use of the text documents created by viewers themselves but not by producers. Furthermore, applying the above approaches to Internet bulletin boards is not effective, because Internet bulletin boards are characterized by their ambiguous structure. Therefore, our challenge is to seek for a technique to extract a kind of patterns out of such ambiguous structured documents.

The important phenomenon our method uses is the synchronization between ongoing TV programs and articles posted to Internet bulletin boards. After our original idea had been proposed in Ref. 9), similar work has been reported by other researchers in Ref. 8). They also report the common existence of similar synchronization. We experimentally confirm the ability of the analysis on this synchronization in Section 5.

A distance education system is proposed in Ref. 13). It also uses synchronization between ongoing video programs and articles posted to bulletin boards. In Ref. 13), time stamps attached to each comment on the bulletin board are used for creating annotations on video lectures. Our proposal is the same in this sense. However, the huge amount the articles used in our study enables the statistical analysis. Over thousands of articles enables the automatic extraction of important keywords.

In terms of practical usage, the successful achievement of our approach will contribute to realizing selective viewing of TV programs in accordance with viewers' attention. Viewers often pay attention to scenes that were not intended to arouse their interest. For example, in the case of drama, background characters often attract more attention from viewers than leading characters. In another case the characters playing villains arouse viewers' sympathy. These facts imply that the intentions of those who produce TV programs do not necessarily meet with the viewers' attention. The viewers' side annotations based on our approach bridge this gap, and enable viewers to select their favorite scenes.

In the experiments reported in Section 5, we extract names of actors as keywords.

3. Statistical Approach for Creating Annotations Based on Internet Bulletin Boards

This section introduces the basic idea for creating annotations out from Internet bulletin boards. Our idea is rooted in the characteristics found in Japan's largest Internet bulletin boards called "Channel 2". In the following subsections, we first explain viewers' common activities that express their attentions. Secondary we explain how these activities explicitly appear on Channel 2, the Japanese largest bulletin board. Then we introduce the basic idea for extracting viewers' attention and creating annotations. After that, a statistical technique is introduced to embody the basic idea, followed by an explanation of graphical output from this technique.

3.1 Viewers Attentions

TV programs often evoke the common excitements among the viewers. For example, viewers watching comedies or action dramas are excited with almost the same scenes, such as scenes of good jokes or thrilling scenes. These excitements often form a kind of communities in which the viewers share their excitements among themselves. For example, viewers come together for drama freaks' communities, and express their excitements with each other. In this situation, we can often find an echo effect. An echo effect is a situation where each viewer says the same meaning word one after another. For example, each viewer mentions the same actor who played the good performance. The echo effect is viewers' activities caused by their desire to share their excitements, and is a very common response among TV viewers. The effects occur not only in real communities like lunch time conversations among office ladies, but also in virtual communities like Internet bulletin boards of drama freaks.

Furthermore, in case of virtual communities, echo effects sometimes occur exactly when the TV programs are on the air. In fact, so called "double screen symptom", meaning viewers manipulate PCs while watching TV, is now prevailing. Two statistical figures support this symptom. One is an article by Nikkei news. According to the article[10], approximately 40 per cent of viewers were in the double screen conditions while watching the Olympic Games in Athens in 2004. Another supporting figure is brought from Hakuhodo, second largest ad agency. Reference 11) shows that 52.9 per cent of viewers have several experiences to use Internet access while watching TV programs. Thus, real time echo effects between double screen players are becoming ordinal phenomena. Reference 15) analyzes how actively the double screeners post articles on ongoing TV programs. Reference 15) also found that the information of articles on ongoing TV programs is far richer than that of non simultaneous articles on past TV programs.

Making use of such real time echo effects on virtual communities, we try to recognize viewers' attention on TV programs.

3.2 Internet Bulletin Boards on TV Dramas

This research focuses on the Japan's largest Internet bulletin board called "Channel 2"[12] as a source of viewers' attention. Channel 2 is a virtual channel which does not exist on terrestrial TV broadcasting in Japan. In fact, Channel 2 is an aggregation of a large number of Internet bulletin boards called "Threads". Different topic is discussed in each Thread. Viewers can selectively participate in their favorite Threads. Out of these Threads, the Threads concerning TV programs are especially extensive and very active. The name of Channel 2 is thought to be derived from such characteristic.

As described earlier, Internet bulletin boards are generally composed of ambiguous description styles, and the same applies to Threads. The dialogues often include jargon that only regular participants can understand their meanings. Their sentences often do not follow proper Japanese grammatical rules. And the rules for making comments are so unrestricted that the dialogues often jump the track from ongoing scenes.

Meanwhile, the Threads concerning TV programs show a typical echo effect as commonly observed in other Internet bulletin boards on TV programs. By this effect, words representing viewers' attention tend to appear with notable frequency as compared with the other words. Furthermore, the frequency can be statistically recognized in Threads, because they are composed of a large volume of comments ranging from 2,000 to 5,000 per one hour of drama program. Thus, our basic idea for recognizing the viewers' attention is to statistically distinguish these words from the unstructured dialogues.

Although the idea above might be effective in

発言番号 発言者ニックネーム 発言時刻 発言者id
920 名無しでいいとも！ New! 04/02/26 22:27 ID:W6Wy+2zi
○○キター―――――――
921 名無しでいいとも！ New! 04/02/26 22:27 ID:gc0dXfab
これ台本がおかしい まるで過失しなければ助かったみたいな感じにしてる
922 名無しでいいとも！ New! 04/02/26 22:27 ID:Zx6nQzS1
○○教授の言うとおり！！！！！！！！！！！
Dialogues in the thread of Channel 2

Seq.No.	Nick name	Time stamp	ID
001	John	22:01	ID: aaaa

" Actor A is not suitable for the main character of this drama."

| 002 | Mary | 22:03 | ID: bbbb |

"I do not think so, Actor A's performance is very good."

The thread of Channel 2 is the repetition of the data consisting of Seq.No., Nick name, Time stamp, ID, and the statement. The explanations of these data items are given below.
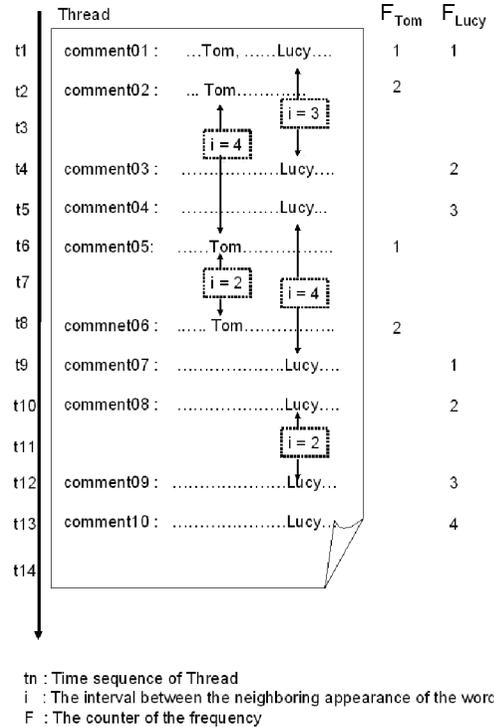
Seq.No.   : The sequence No. of the statement
Nick name : Nick name of the participant
Time stamp : The time when the statement was made by the participant
ID   : ID of the participant

Fig. 1 Live Thread in Channel 2.

tn : Time sequence of Thread
i : The interval between the neighboring appearance of the words
F : The counter of the frequency

Fig. 2 "Extracting symbolic words".

terms of recognizing the statistical regularities out of unstructured dialogues, it is not sufficient for creating the annotations, because we still do not know corresponding scene that the extracted words indicate. We need the information that enables us to recognize the linkage between these words and corresponding scenes in the TV program. To overcome this issue, we make use of the time stamps attached to each comment in the Threads. **Figure 1** shows the layout of a Thread. Each comment is attached with the time stamp indicating when the participant make their comments. When viewers participate in the Threads, they make their comments exactly when their favorite scenes come out. Naturally, these time stamps represent the scenes that each comment mentions about. Thus, recognizing symbolic words attached with these time stamps leads to extracting information of viewers' attention that act as viewers' side annotations. Hereafter, we call the words in echo effects symbolic words. Following subsections explain a statistical technique for extracting symbolic words.

3.3 The Method to Extract Symbolic Words

Symbolic words frequently appear in the Thread when viewers' favorite scene comes out. To extract such symbolic words from Threads, we need some methods which measure frequency of words. Term frequency and inverse document frequency, i.e., TF/IDF, is a well-known method which measure the frequency of words. Although the method is adequate to measure the frequency of words in a given document, we need a slightly different method. In our case, the frequency should be counted by each period of echo effect,which means that the corresponding segment for each echo effect in the Thread should be recognized as the separate documents. TF/IDF does not provide with the way for the segmentation.
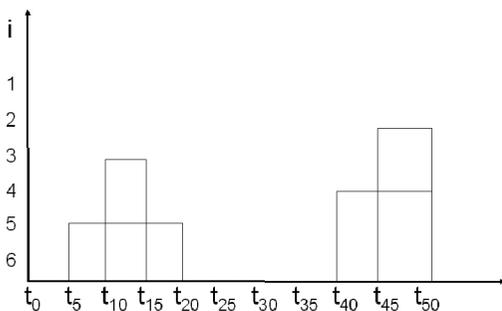
To cope with this issue, we propose a method, which can recognize temporally fluctuating segments. While an echo effect occurs, corresponding symbolic word reappears consecutively within considerably short intervals. Making use of this characteristic, our rule examines the intervals between neighboring appearances. Only if the observed intervals are short enough, the appearance is counted as the frequency. The words showing outstanding frequency under this rule is recognized as symbolic words. **Figure 2** shows the idea. The vertical axis represents the time sequence of the Thread. In this example, the threshold of the interval is initialized as 3, and the threshold of the frequency is initialized as 3. Given these values, frequency of the word Tom is reset at t5 to 1 because the interval between t1 and t5 is greater than 3. After that the frequency is counted up to 2 at t6. Tom is not recognized as a symbolic word, because it never surpasses its threshold of the frequency. The word Lucy is counted

up to 3 at t5, because it consecutively appears with shorter intervals than 3. And again Lucy is counted up to 4 at t13 after the frequency is reset to 1 at t9. In this example, Lucy is recognized as a symbolic word twice, i.e., from t1 to t5, and from t9 to t13. Although the words are the same, they are different symbolic words in the sense that they represent different echo effects.

3.4 Attention Graph, Graphical Structure of Viewers' Side Annotation

In the event of echo effects, the more enthusiastically viewers pay attention to a scene, the more frequently the symbolic words appear. The appearances with shorter intervals indicate higher degree of enthusiasm. To measure the degree of enthusiasm, we apply the extraction method above by varying the value of the interval. Narrower interval results in extracting symbolic words representing stronger enthusiasm, and vice versa.

Figure 3 shows the example. Note that this figure is the expansion of Fig. 2, in terms of adding the variable intervals (i.e., degree of enthusiasm) on the vertical axis. The time sequence of the Thread is on the horizontal axis. In this example, two symbolic words are observed. One of them starts to appear at t5 and ends at t20, indicating the interval is less than 5. Inside this period, there is a graph indicating a narrower interval, 3 from t10 to t15 which indicates stronger echo effects. Multiple extractions with different values of intervals enable analysis of enthusiasm. Similarly, another symbolic word starts to appear at t40 and ends at t50 with the interval of less than 4. This indicates the stronger echo effect than the first one. Again, there is the period inside where the symbolic words appear more frequently at

the intervals less than 2.

As the example shows, multiple executions of the extraction rule construct mountaneous graph structure. The height of the graph indicates the degree of enthusiasm. On the other hand, the horizontal position acts as the indexes to the corresponding scenes that viewers actually keep their attentions. Starting time and ending time of the position are synchronized with ongoing TV program. We call this graph "attention graph". Attention graph acts as viewers' side annotations representing the scenes that viewers pay attention to as well as their degrees of enthusiasm. It also has the information of symbolic words.

4. TV Drama Annotating System

4.1 System Configuration

Figure 4 illustrates the system configuration of a prototype TV drama annotating system. The component ① records TV drama and produces an mpeg4 video stream.

While the recording component ① is recording TV drama, the corresponding Thread is downloaded and a morphological analysis is applied to extract words from the Thread (②). The preprocessing component ② also deletes stop words and consolidates different expressions of the same meaning words, such as jargons or nicknames of actors into single words. For the consolidation, the preprocessing component ② makes use of dictionary that defines different expression of the same meaning words

In order to construct the dictionary, we used the information published on the official web site of the drama, and manually refined its contents. Since most of the official web sites of
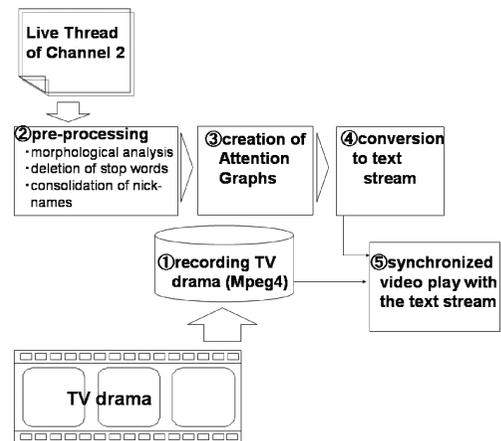

t : The time sequence of the Thread
i : The neighboring time interval

Fig. 3   Attention graph.

Fig. 4   System configuration.

the dramas have lists of the actors' full names, we used them as the primary dictionary. From this information, we first break down an actor's full name into a first name and a last name and registering them separately. For example, if the actor's name is Takuya Kimura, we just register both Takuya and Kimura separately as the aliases of Takuya Kimura. We also added actor's popular nick names if they have. In the case of Takuya Kimura, we added KimuTaku as the alias of Takuya Kimura.

In order to consolidate jargons, we made use of "Ni-channel jisyo"[14]. Although the dictionary we used is simple, it is enough to help our main procedures which statistically analyzes keywords in the articles.

In the graph creation component ③, we implement the algorithm which embodies the extraction method of attention graph described in Section 3 (See next subsection for the details.). Component ④ just converts output of ③ into text stream data. Finally, component ⑤ merges the video stream from ① and the text stream ④ and enables to view the TV programs synchronized with attention graphs.

4.2 Algorithm for Generating Attention Graphs

The algorithm implemented on graph creation component is described on **Fig. 5**. It inputs two files. "TimeStampedWordsFile", the output file from the pre-processing component, has the list of morpholized words from Live Threads. Each word in this file has time stamp derived from the Live Thread, thus all the words are sorted out by the time sequence. Another file is KeyWordFile. This file includes the same words as TimeStampedWordsFile. The differences are that the words in this file are attached with threshold on interval, and each word is uniquely registered.

The algorithm first reads a word from KeyWordFile then, searches for the same word in TimeStampedWordsFile. Every time when the word is found, the algorithm calculates the interval between the time stamp of the word and the one of the word previously found. If the interval is less than the threshold, the algorithm tentatively stores the word with its time stamp in WkAttnGraphRec then increment the counter. On the other hand, if the interval surpasses the threshold, the algorithm checks if the counter of the word in WkAttnGraphRec surpasses the threshold of the frequency. If it is true, stored data in WkAttnGraphRec is

```
Main Procedure  createAttentionGraphs
Input
    TimeStampedWordsFile : Word list with time stamps
    KeyWordFile : Words list  with the threshold "Th2_w"

Output
    AttnGraphFile ; Attention Graph data comprised of Symbolic
                         Words , durations, Attention Levels
Var
    WkAttnGraphRec ; Work record tentatively storing
                         information for Attention Graph

Begin
    Th1 = threshold of consecutive appearance of w

    while (Read a word and its Th2_w from KeyWordFile) {
     // Every time when a word is picked out from KeyWordFile,
       reading procedure of TimeStampedWordFile starts over
       from the first record.
       while ( read a word from TimeStampedWordsFile ){
           for ( decrement Th2w until it reaches to 1  ){

             if ( the word from KeyWordFile matches to the
                word from the word from TimeStampedWordsFile
                &  interval between this match and  last match is
                less than Th2_w )  {
                  update  ending time of Attention Graph in
                    WkAttnGraphRec, and increment the number
                    of matches
             }

             else if(the word from KeyWordFile matches to the
                word from the word from TimeStampedWordsFile
                &  interval between this  match and  last match is
                greater than Th2w )  {
                 if ( The number of  matches greater than Th1) {
                     write WkAttnGraphRec  to AttnGraphFile
                 }
                 set  time stamp from TimeStampedWordsFile
                    to  WkAttnGraphRec  as the starting time of
                    Attention Graph
                 set  a word from TimeStampedWordsFile to
                    WkAttnGraphRec as a Symbolic Word
             }
          }
       }
       Write all the data remaining in WkAttnGraphRec to
            AttnGraphFile when reading procedure of
            TimeStampedWordsFile reaches to the end
    }
End
```

Fig. 5 Algorithm for generating attention graphs.

recognized as symbolic words, and written in AttnGraphFile. If it is false, stored data is cleared out. By decrementing the threshold from given threshold down to 1, this algorithm extracts every periods which represent different degrees of viewers' enthusiasms. Here the periods share the same symbolic words. With higher threshold, periods of higher enthusiasms are extracted. With a smaller threshold, periods of lower enthusiasms are extracted. After that, the next same word is searched in TimeStampedWordsFile. If it is found, the threshold is reset back to the given value. Then the algorithm again extracts the periods of viewers' attention. This procedure is repeated by decrementing the threshold until it reaches to 1.

When TimeStampedWordsFile reaches to its end, the algorithm read the next word from KeyWordFile, and start over all the procedure above from the first record in TimeStampedWordsFile to the end. Thus, all the data for attention graphs are written in AttnGraphFile.

5. Evaluations

To evaluate the performance of the attention graphs, we have executed three types of tests. In this section, we start with the synchronicity test, which is to verify that the positions of the attention graphs act as the indexes to the scenes of dramas. Then, next test is to verify that the positions of attention graphs are coincident with the ones viewers actually pay attention to. Final test is to show that the heights of the attention graphs coincide with the degrees of viewers' enthusiasm.

In these tests, the candidates for symbolic words are limited to the actors' names. This is based on the preliminary study we performed. We invited actual viewers and showed them the drama program to ask them which scenes they were interested in. During the test, we also asked them about what information contributed to their interest on each scene. For this purpose, we prepared 6 categories, i.e., actors' shape, actors' performance, scenario, acting scripts, setting of the scene, and gadgetries as the candidates of the contributing factors. Then, we asked the viewers to prioritize these categories for each scene of interest.

Based on the test, we carried out principal component analysis, and found that actors are the most contributing factor of viewers' interests. **Table 1** shows the results in detail. The principal eigenvector on the first column shows relatively large weights on actors' shape, and actors' performance. Thus, this eigenvector can be interpreted as the contributing factor concerning actors. In addition to that, weight of the contribution of the eigenvector is 31.1, which is far larger than the contribution weight of the second eigenvector, 17.8. Since actors' names are the most important words for the

Table 1 The contributing factor for viewers' interests.

Eigenvectors	1st	2nd	3rd	4th	5th	6th
Actors' shape	0.6	-0.1	-0.1	0.1	-0.4	-0.7
Actors' performance	0.6	-0.3	-0.1	0.1	-0.4	0.7
Scenario	0.2	0.7	0.3	0.6	-0.1	0.1
Acting scripts	0.4	0.1	-0.5	0.1	0.7	0.0
Settings of the scene	0.3	0.4	0.4	-0.8	0.1	0.1
Gadgetary	0.2	-0.5	0.7	0.3	0.4	-0.1
Eigenvalue	1.9	1.1	1.1	0.8	0.7	0.5
Contributing weight	31.1	17.8	17.7	13.8	11.2	8.5

annotations in terms of mostly reflecting viewers' interests, we only select actors' name in the experiments.

5.1 Synchronicity Test of Attention Graphs

If the positions of attention graphs correctly reflect the times at which the viewers' attention is kept alive, each attention graph should appear exactly when corresponding scene appears. For example, the attention graph labeled by the symbolic word Tom starts to appear at 24 minutes 10 seconds after the start of the show and disappear at 25 minutes 25 seconds, Tom should actually appear in the scene. For this test, we have taken 20 minutes of drama and corresponding Thread. In this test the candidates for the symbolic words are simply limited to actors' names. Under these conditions, we executed the system shown in Fig. 4 and compared the video stream with the text stream to see if each attention graph appears exactly when the corresponding actors appear in the scene.

The system created 30 attention graphs from a 20-minute sequence of drama, and we verified 23 out of them are synchronized with the actors who actually appear in the scenes. Remaining 7 are all errors. 5 attention graphs out of these errors found to appear while the video stream shows commercial messages. Participants in the Threads keep chatting during commercial message breaks. They tend to review the most attractive scenes before the commercial messages. The reason of 5 errors are attributed to these kind of participants' behavior. If we eliminate these errors created during commercial messages, attention graphs resulted in 23 matches out of 25, which is satisfying figure equivalent to 92 percent matching ratio.

5.2 Coincidence of Attention Graphs with Viewers' Attentions

The second and third tests are to verify that attention graphs coincide with viewers' attention and their degree of enthusiasm. For this purpose, we have invited 27 viewers, showing them a 20-minute sequence of drama, and asking them manually construct attention graphs. Then, we compared manually constructed attention graphs with the attention graphs created by our system. The second test is to check if the positions (i.e., starting time and ending time) of both attention graphs coincide with each other. **Figure 6** illustrates the results. Horizontal columns represent the actors'

Fig. 6 The positions of attention graphs.

Table 2 The coefficients of correlations.

Actor A	0.7
Actor B	0.7
Actor C	0.9
Actor D	0.9
Actor E	0.9
Actor F	0.9
Actor G	0.6

statistical correlations, we labeled the positions with shadowy bars "1", and labeled the positions without any shadowy bars "0". In this way, we have calculated the coefficients of correlations between corresponding shadowy bars. The result of the calculation is shown in **Table 2**. Most of the correlations are ranging from 0.7 to 0.9. Thus, the result confirms that the positions of the automated attention graphs correctly reflect where the echo effects actually occur.

5.3 Coincidence of the Degrees of Enthusiastic

The third test is to verify that the attention graphs correctly reflect the degree of viewers' enthusiasm. For this purpose, we additionally asked invited viewers to give their subjective degrees of their enthusiastic to each manually constructed attention graph. The degrees are expressed in the figure ranging from 1 to 5. 1 means the lowest degree. Because we found the viewers felt difficulties to express their degrees of enthusiasm by every tiny intervals, such as 10 seconds, we just asked them to point out the most impressive time in each attention graph and label the degree.

Again **Fig. 7** shows the comparison of manual attention graphs with the automated ones. While Fig. 6 shows the comparison of the positions of the attention graphs, Fig. 7 shows the comparison of the degree of enthusiasm as well. The horizontal axis represents the time sequence, and the vertical axis represents the degree. The graphs are separately depicted by each actor's name, A to G. The dark shadowy sequential line graphs represent the automated attention graphs, and the light shadowy ones represent the manual attention graphs.

Underlying assumption to depict the light shadowy graphs is that the degree of enthusiasm given by the viewers would remain unchanged throughout the duration of the attention graph. If multiple viewers labeled different points with different degrees for the same attention graph, the degrees are assumed to decline or rise linearly between the points.
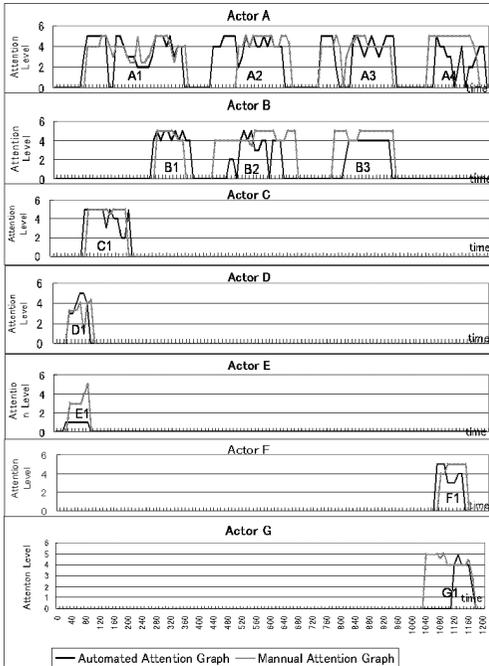
names labeled to each attention graph. Each column is divided into two sub-categories. Under these sub-categories, shadowy bars appear along with the vertical axis, which represents the time sliced with 10-second intervals. The dark shadowy bars under the left sub-categories represent the positions of the automated attention graphs. The light shadowy bars under the right sub-categories represent the positions of the manual attention graphs. If the positions of automated attention graphs correctly reflect where the echo effects occur, the dark shadowy bars should appear at the same time with the light shadowy bars.

As shown in this figure, most of the corresponding shadowy bars appear approximately at the same times. To evaluate this result with

Fig. 7   Comparison of attention levels.

In principle, Fig. 7 shows the coincidence of the heights between both the attention graphs, which indicates automated attention graphs correctly reflect viewers' enthusiasm. Several incoincidences are found between the graphs. Typical examples are the graphs labeled as A2 and B2. The positions of both of A2 and B2 are corresponding to the scenes where Actor A seriously argues with Actor B. Most of the viewers are thought to feel enthusiastic to the argues themselves, but not the actors. Actors' names just act as the symbols mark of the arguing scenes in this case. And participants in the Thread seem to symbolize the arguing scenes by Actor A, while the invited viewers seem to symbolize them by Actor B. Similarly, the graphs D1 and E1 represent the arguing scene between Actor D and Actor E.

6.   Conclusions

In this paper, we have proposed a method for creating viewers' side annotations to TV dramas from the viewers' dialogues on Internet bulletin boards. In general, dialogues on Internet communities are often expressed in a poor grammatical manner, preventing the previous methods from being applied effectively. In order to overcome this issue, we proposed a statistical method which enables to count the

frequency separately by each temporally fluctuating segment. The results from applying the method can be expressed in the form of attention graphs. The performance tests of attention graphs show that:

(1)   The positions of attention graphs act as indexes to the scenes that viewers pay attention to.

(2)   The heights of attention graphs generally reflect the degree of viewers' enthusiasm to each scene.

In summary, attention graphs succeeded in expressing two properties belonging to viewers' attention. One is the positions of the attention, and another one is the degree of enthusiasm. Although attention graphs tell us about what the viewers pay attention to by symbolic words, they do not tell us how the viewers feel. For example, in case of the symbolic word Tom, we can not tell that the underling feeling is empathy for him or hate for him. Internet bulletin boards are filled with viewers' feelings, there are some possibilities to extract these kinds of viewers feelings to each scene. Our next issue will be to seek for the method which enables recognition of viewers' feelings.

References

1) Uehara, H. and Yoshida, K.: Annotating TV Drama based on Viewer Dialogue, *Proc. SAINT2005*, pp334–340 (2005).

2) Okamoto, M., Kamahara, J., Ueda, K., Shimojo, S. and Miyahara, H.: Design and Implementation of Automatic Sports Digest Recomposition System with Scenario Template, *The Institute of Electronics, Information and Communication Engineers Transactions*, Vol.J85-B, No.8, pp.1269–1276 (Aug. 2002).

3) Hamada, R., Ide, I., Sakai, S. and Tanaka, H.: Structural Analysis of Cooking Preparation Steps, *The Institute of Electronics, Information and Communication Engineers, Transactions*, Vol.J85-D-II, No.1, pp.79–89 (Jan. 2002)

4) Hamada, R., Ide, I., Sakai, S. and Tanaka, H.: Associating cooking video with related textbook, *Proc. ACM Multimedia 2000 Workshops*, pp.237–241 (2000).

5) Ide, I., Yamamoto, K., Hamada, R. and Tanaka, H.: An automatic video indexing method based on shot classification, *Systems and Computers in Japan*, Vol.32, No.9, pp.32–41 (Aug. 2001).

6) Yaginuma, Y., Izumi, N. and Sakaguchi, M.: A Proposal of Video Editing Method Using Synchronized Scenario Document, *The Insti-*

tute of Electronics, Information and Communication Engineers, Transactions, Vol.J79-D-II, No.4, pp.547–558 (Apr. 1996).

7) Moriyama, T. and Sakauchi, M.: Video summarization based on the psychological content in the track structure, *Proc. ACM Multimedia 2000 Workshops*, pp.191–194 (Nov. 2000).

8) Miyamori, H., Nakamura, S. and Tanaka, K.: Automatic indexing of broadcasting content using its live chat on the web *Proc. PRMU Workshops*, Vol.104, No.668, pp.43–48 (Feb. 2005).

9) Uehara, H. and Yoshida, K.: Scripting TV Drama based on Viewer Dialogue — Analysis of viewers' Attention on an Internet Bulletin Board, *Proc. PRMU Workshops*, Vol.104, No.369, pp.25–30 (Oct. 2004).

10) Internet ads make rapid progress. http://homepage1.nifty.com/o-terue/nikkei/2005/idex02.html

11) Growth of double screen life style by prevailing broadbands. http://www.watch.impress.co.jp/internet/www/article/2002/1011/hakuhodo.htm

12) Channel 2. http://www.2ch.net/

13) Haga, H.: Combining video and bulletin board systems in distance education systems, *Internet for Higher Education*, Vol.5, pp.119–129, Elsevier Science (2002).

14) Ni-channel Jisyo. http://www.geocities.co.jp/Playtown-Denei/3081/dic/2ch/

15) Matsumura, N., Miura, A., Shibanai, Y., Ohsawa, Y. and Ishizuka, M.: The Dynamism of 2 channel, Vol.45, No.3, pp.1053–1061 (Mar. 2004).

Hiroshi Uehara now enrolled in Graduate School of Business Science, University of Tsukuba. At NTT DoCoMo Ltd., his research focuses on mobile Internet applications supporting TV broadcasting services for mobile handsets. He is currently a member of the IEEE.

Kenichi Yoshida received his Ph.D. from Osaka University in '92. In '80, he joined Hitachi Ltd, and is working for University of Tsukuba from '02. His current research interest includes application of internet and application of machine learning techniques.