

Efficient Tracking of News Topics Based on Chronological Semantic Structures in a Large-Scale News Video Archive*

Ichiro IDE^{†,††a)}, Senior Member, Tomoyoshi KINOSHITA^{†††}, Tomokazu TAKAHASHI^{††††}, Hiroshi MO^{††}, Norio KATAYAMA^{††}, Shin'ichi SATOH^{††}, Members, and Hiroshi MURASE[†], Fellow

SUMMARY Recent advance in digital storage technology has enabled us to archive a large volume of video data. Thanks to this trend, we have archived more than 1,800 hours of video data from a daily Japanese news show in the last ten years. When considering the effective use of such a large news video archive, we assumed that analysis of its chronological and semantic structure becomes important. We also consider that providing the users with the development of news topics is more important to help their understanding of current affairs, rather than providing a list of relevant news stories as in most of the current news video retrieval systems. Therefore, in this paper, we propose a structuring method for a news video archive, together with an interface that visualizes the structure, so that users could track the development of news topics according to their interest, efficiently. The proposed news video structure, namely the “topic thread structure”, is obtained as a result of an analysis of the chronological and semantic relation between news stories. Meanwhile, the proposed interface, namely “mediaWalker II”, allows users to track the development of news topics along the topic thread structure, and at the same time watch the video footage corresponding to each news story. Analyses on the topic thread structures obtained by applying the proposed method to actual news video footages revealed interesting and comprehensible relations between news topics in the real world. At the same time, analyses on their size quantified the efficiency of tracking a user’s topic-of-interest based on the proposed topic thread structure. We consider this as a first step towards facilitating video authoring by users based on existing contents in a large-scale news video archive.

key words: news video, video archive, semantic structure, topic tracking, topic threads

1. Introduction

Recent increase of commercially available digital storage capacity has allowed us to archive a large volume of video data. Among various types of video, especially broadcast news video is a rich source of information concerning the human society, which is worth archiving and subsequently, retrieving and reusing. Thus, starting from the

pioneering works in the Carnegie Mellon University’s Informedia News-on-Demand project [3], many researchers have worked on indexing, retrieving, summarizing news video, and so on.

However, most of the works attempt to make use of the video data in an archive as they are. Such technology is indeed necessary, but now that a news video archive could easily house video footages from years of news shows covering the cause and result of various topics, simply presenting a list of news stories that match a query based on the indexing and retrieving technologies, or simply showing a summary of a video footage from one news story is insufficient. It is the time that we started handling the data according to their semantic relations together with their chronological nature, so that users could retrieve and understand the development of a news topic concerning their query, or reuse them to visually “tell stories” based on their own points-of-view based on the facts.

In order to support users “tell stories”, it is necessary to allow them to easily gather and select relevant video footages that could be used as source materials, without being forced to browse through a long list of video footages which may contain redundant or noisy contents. There are several works on video authoring that aims to generate a new video according to the user’s intention by combining existing video materials such as Bocconi et al.’s “Vox Populi” system [4] that generates a documentary video by combining detailedly annotated monologue video clips, and Shrestha et al.’s system on socially gathered concert video clips [5]. Meanwhile, the authoring of a video considering a large-scale news archive as a source of materials has not been challenged before. So in this paper, as the first step towards facilitating video authoring based on existing contents in a news archive, we propose a structuring method of the archive, together with an interface that visualizes the structure.

In the following, after introducing definitions of terms used throughout the paper in Sect. 2, we first present a news video structuring method based on the chronological semantic relations between news stories (“topic threading”) in Sects. 3 and 4, and next introduce in Sect. 5, an interface based on the structure (“mediaWalker II”) which facilitates users to track news topics along the timeline, and at the same time, allows them to efficiently choose video sources so that after post-editing, they could visually tell their own stories based on the video footages.

Manuscript received May 2, 2011.

Manuscript revised August 31, 2011.

[†]The authors are with the Graduate School of Information Science, Nagoya University, Nagoya-shi, 464–8601 Japan.

^{††}The authors are with the National Institute of Informatics, Research Organization of Information and Systems, Tokyo, 101–8430 Japan.

^{†††}The author is with NetCOMPASS, Ltd., Tokyo, 104–0033 Japan.

^{††††}The author is with the Faculty of Economics and Information, Gifu Shotoku Gakuen University, Gifu-shi, 500–8288 Japan.

*This paper is a revision of the paper presented at the Second Annual Summit and Conference of the Asia-Pacific Signal and Information Processing Association [1] and [2].

a) E-mail: ide@is.nagoya-u.ac.jp

DOI: 10.1587/transinf.E95.D.1288

Section 6 introduces related works in both the structuring and the visualization of news video, and Sect. 7 concludes the paper.

2. Definition of Terms

Some terms specific to news are used throughout this paper according to the following definitions. First, the following three terms follow the definitions in the Topic Detection and Tracking (TDT) workshop series organized by NIST [6].

- **Event:** Some incident that occurred at some specific time and place along with all necessary preconditions and unavoidable consequences.
- **Story:** A topically cohesive segment of news that includes two or more declarative independent clauses about a single event.
- **Topic:** A seminal event or activity, along with all directly related events and activities.

Next, the following three terms are defined by the authors.

- **Topic thread:** A sequence of related stories chained chronologically. It may contain several topics.
- **Topic thread structure:** A directed graph composed of topic threads originating from a specified story.
- **Topic cluster:** A topically cohesive set of neighboring stories in the topic thread structure.

3. Analyzing the Chronological Semantic Structure of News Stories

3.1 Overview

Figure 1 (a) shows the general structure of a news video archive; It is composed of videos recorded from daily news shows where each of them consists of several stories. Story segmentation is one of the oldest topic and a basic technology in the news contents analysis field. The retrieval of segmented stories has also been approached from both keyword querying and content-based image retrieval. Especially, image-based retrieval has been widely studied recently in the TRECVID[†] community.

These are indeed essential functions for a news video archive, but we assume that a user would generally seek for the knowledge on how a topic developed along the time as shown in Fig. 1 (b), rather than details on individual stories.

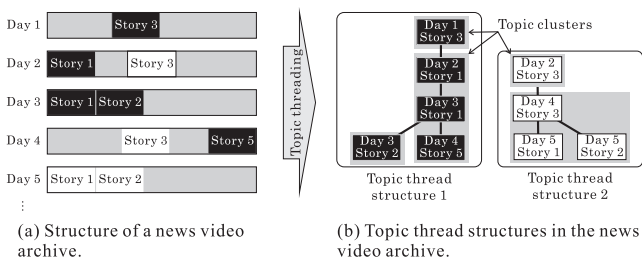


Fig. 1 Semantic structure in a news video archive.

In other words, each story provides information mainly on the so-called 4Ws (When, Where, Who, and What) of an event, while we assumed that a user would more likely prefer to understand the 1W1H (Why and How), which requires the browsing of the development of topics along the timeline.

Under this assumption, we introduce a method that chains individual stories into a “topic thread structure” according to their chronological and semantic relations (Fig. 1 (b)). It represents local relations as directed edges, and at the same time a global trend of topics as a directed graph. Figure 2 shows an example of a large topic thread structure obtained by the proposed method. We can see that the structure not only chains obviously related topics but also topics without obvious relations, although they are actually related in the real world. The proposed method should help a user understand the actualities of such relations.

The following process is required to obtain a topic thread structure, and their details are introduced in the following Sections.

1. **Story segmentation:** Each day’s news video is segmented into stories.
2. **Topic threading:** Topic threads are extracted based on chronological and semantic relations between all pairs of stories.
3. **Extraction of topic clusters:** Topic clusters are extracted within a topic thread structure.

3.2 Story Segmentation

Many story segmentation methods for news video has been proposed by various groups [8], and most of them make use of visual features even in early works in the field [9], [10]. Regardless of the existence of these works, since we work on an archive that covers a long period of time, in order to avoid dealing with the occasional change of studio designs and editing policies in the show, we decided to make use of only the closed-caption (an audio transcript provided digitally from the broadcasters) for story segmentation^{††}. The following steps were applied to the closed-caption obtained from each news show. Details of the method are found in our previous publication [11], [12].

1. Apply morphological analysis^{†††} to each sentence of the closed-caption text, and extract noun compounds.
2. Classify the noun compounds into four semantic attributes; general, personal, locational/organizational,

[†]See “TREC Video Retrieval Evaluation conference series,” <http://trecvid.nist.gov/> together with Smeaton et al.’s publication [7].

^{††}The closed-caption used in the following experiments was automatically synchronized to the audio channel at the time of archiving.

^{†††}A Japanese morphological analyzer JUMAN developed at Kyoto University was used. It is available for download from <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/juman-e.html>.

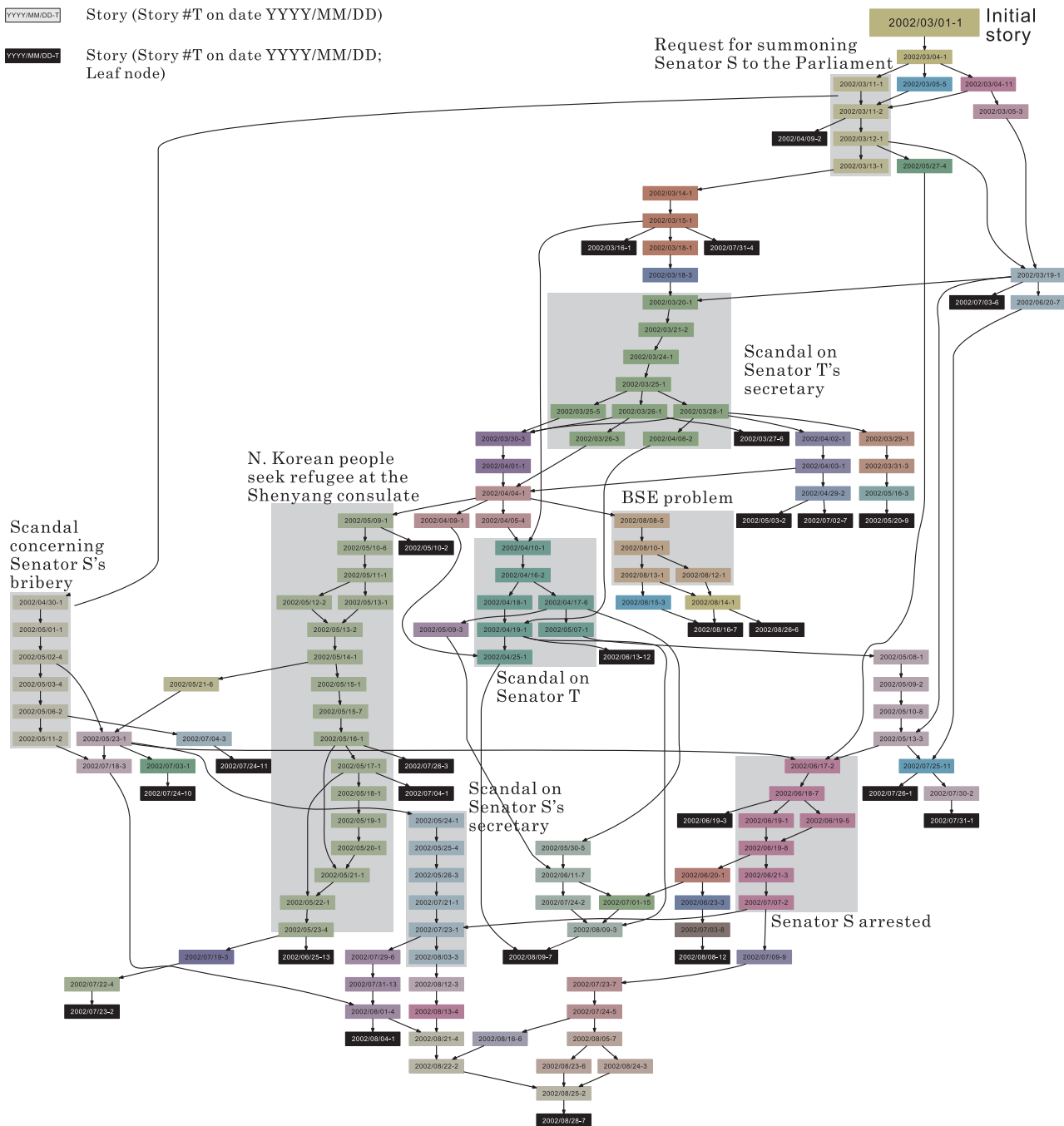


Fig. 2 Example of a topic thread structure originating from the first story on March 1, 2002 with a search period of $d = 180$ days and a relation threshold $\theta_{trk} = 0.40$ and a clustering $\theta_{cls} = 0.40$. It took 2.52 secs. to obtain this structure. Some large topic clusters are labeled manually by the authors to demonstrate the transition of the actual topics.

- or temporal, based on our previous work [13], and create a term frequency vector for each attribute.
- 3. At each sentence boundary, concatenate w preceding and w succeeding term frequency vectors, and compare them in each semantic attribute. The window size w is changed from 1 to N_w , and the maximum similarity of the vectors measured in cosine measure is selected.
- 4. Combine the similarity in each semantic attribute as a weighted sum. In order to reflect the general dif-

- ference in the importance of terms in each semantic attribute, an empirically obtained weight of (general, personal, locational/organizational, temporal) = (0.23, 0.21, 0.48, 0.08) was applied. If the combined similarity falls below a threshold θ_{seg} , a story boundary is detected there.
- 5. After applying steps 1.–4. to the entire closed-caption text, create a term frequency vector for each story, and rejoin neighboring stories if the similarity between

them exceeds a threshold θ_{cat} . This step is recursively applied until no more rejoining occurs.

In the following experiments, the parameters were set empirically as $N_w = 10$, $\theta_{seg} = 0.17$, $\theta_{cat} = 0.13$, according to the balance of precision and recall in a preliminary experiment [12]. An evaluation on fourteen news shows with 130 ground truth stories showed a segmentation ability of 95.4% recall and 90.5% precision when an error of plus/minus one sentence was allowed. Since there tend to be redundant comments near story boundaries, we consider that allowing an error of plus/minus one sentence is reasonable.

3.3 Topic Threading

Next, a topic thread structure originating from a specified story is created by the topic threading method. The purpose of creating a topic thread structure is to chain related stories along the timeline in order to provide a user with paths, namely “topic threads,” to follow the development of topics.

The simplest solution for this task is to create a tree structure by recursively expanding related stories considering the chronological order (Fig. 3 (a)). However, when a news video archive covers a long period, this solution results in creating a tree structure with numerous branches at each node (story), and also numerous duplicates of the same story appearing in nodes all over the tree which makes the structure redundant. Tracking news topics story by story along such a structure requires watching a large number of video footages at each node to select an interesting story in the succeeding nodes, besides an occasional encounter with a video footage that has already been watched in a previous node due to duplicates.

The proposed topic threading method is designed so that it should avoid these problems. First, it unifies all the nodes with the same story to one node that appears in the deepest hierarchy of the tree. This eliminates the redundancy of the structure, so a user will not encounter the same video more than once during the tracking. Next, upon unification, branches are removed or reconnected under certain conditions. This minimizes the number of branches at each node, so the time needed to watch video footages at each

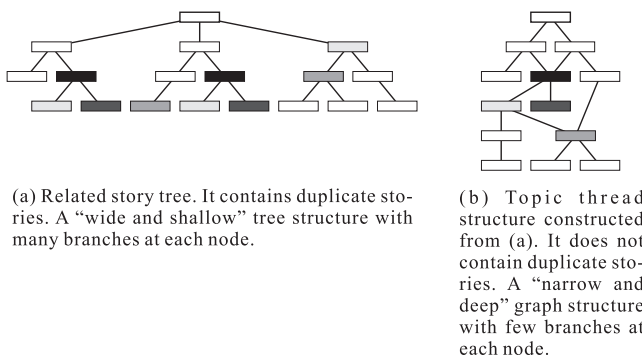


Fig. 3 Representation of the chronological semantic relations between news stories: from (a) a related story tree to (b) a topic thread structure.

node for the selection of the path to follow next, is reduced to the minimum. In the end, we obtain a directed graph structure, namely a “topic thread structure” (Fig. 3 (b)). Details of the topic threading procedure is as follows.

1. **Evaluation of the relation between stories:** The relation between two stories are measured by the cosine distance of their term frequency vectors. The vectors are created and compared in a similar way as in the story segmentation algorithm introduced in Sect. 3.2. A pair of stories considered as related has a relation larger than a threshold θ_{trk} .
2. **Creation of a story relation tree (Fig. 3 (a)):** Create a tree T_r originating from an initial story S_0 by recursively expanding story relations under the following conditions:
 - a. A child represents a story related to, and newer than the story its parent does.
 - b. Among the siblings, the younger represents a newer story.
3. **Unification of nodes with the same story (Fig. 4):** For all the sub-trees $T_s(i)$ in T_r , when there exists an equivalent sub-tree $T_s(j)$ branching with itself from an elder node, unify $T_s(i)$ to $T_s(j)$, and apply either of the following operations to the branch to the root node of $T_s(i)$:
 - a. **Removal:** When $T_s(j)$ is a descendant of $T_i(i)$'s elder sibling, remove the branch.
 - b. **Reconnection:** When $T_s(j)$ is a descendant of $T_s(i)$'s ancestor (except its parent), reconnect the branch to the root node of $T_s(j)$.

Note that due to the removal of an edge (branch) in step 3.-a, we can guarantee that all the siblings are independent; they are not related mutually. This feature makes the selection among the child nodes by a user meaningful during the topic tracking process in a browsing interface, in the sense that the user would generally not need to eventually come back to a child node after tracking down a topic thread originating from its siblings, because they contain different topics than the previously selected topic thread.

The above algorithm is for creating a topic thread structure in the future direction. To obtain a structure in the past direction, “newer” should be substituted with “older” in the algorithm.

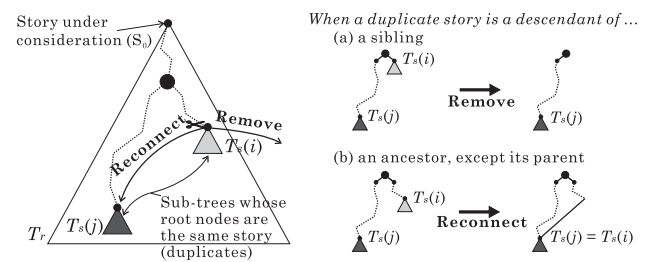


Fig. 4 Identical node unification operations in the topic threading algorithm.

3.4 Extraction of Topic Clusters

Since the topic thread structure is constructed by recursively expanding a story relation tree based only on local relations between stories, it may contain several locally related, but globally different topics that gradually developed along the time. In order to reveal such sub-structures, topic clusters which contain homogeneous stories along the topic thread structure are extracted. Note that the topic clusters are overlapped with the topic thread structure, and are not obtained by clustering all the stories that compose the topic thread structure, as it is generally the case in other works, which will result in ignoring the original topic thread structure. Details of the topic cluster extraction procedure is as follows.

1. Set story S_0 as the cluster center ($C_0 = S_0$) and the story-in-focus ($S = S_0$).
2. Let the children of story S be $S_c(j)$ ($j = 1, \dots, J$). When the relation between stories C_0 and $S_c(j)$ is larger than θ_{cls} , set $S_c(j)$ as a new cluster center ($C_0 = S_c(j)$).
3. Apply step 2. recursively by scanning the entire topic thread structure, shifting S one story after another.
4. Label stories between each topic cluster and the next ones with a same cluster number.

3.5 Example of a Topic Thread Structure

Let's see what a topic thread structure represents by manually analyzing the structure.

Figure 5 (a) shows an example of a topic thread structure obtained by the proposed method. Figure 5 (b) shows a degenerated representation of the topic thread structure by topic clusters, with Table 1 summarizing their actual contents which were manually analyzed and annotated.

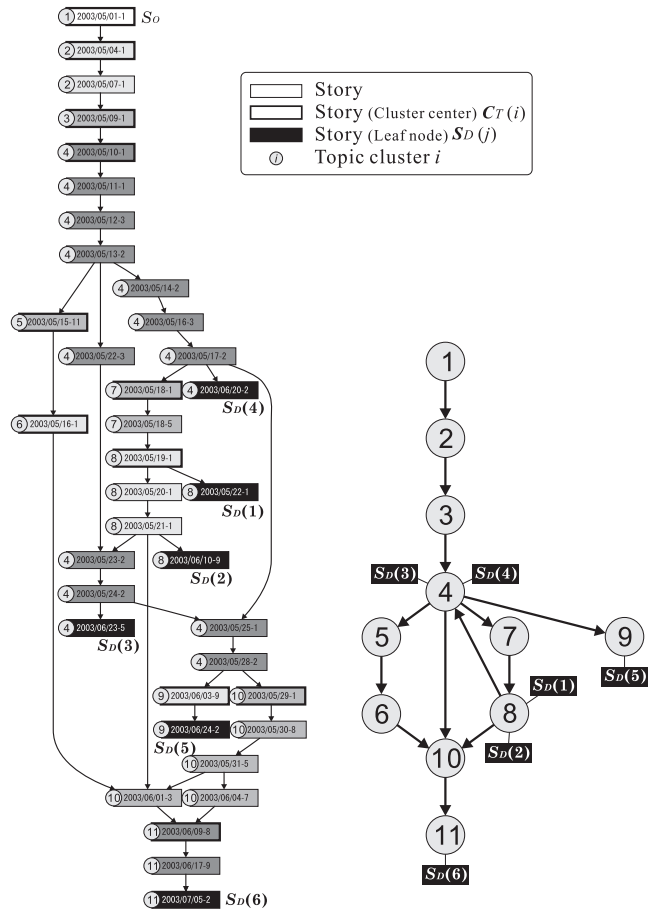
From this structure, for example, we can choose a topic thread $\{C_T(1), C_T(2), C_T(3), C_T(4), C_T(10), C_T(11)\}$, where we can follow the main-stream development of the SARS (Severe Acute Respiratory Syndrome) outbreak in 2003. Likewise, "side stories" such as a partial topic thread $\{C_T(4), C_T(5), C_T(6), C_T(10)\}$, lets us follow the development of the outbreak inside China, $\{C_T(7), C_T(8)\}$ on a false alarm in Japan.

As seen from this example, the topic thread structure represents the various topics that occurred as a consequence of the initial story under consideration. At the same time, users can and needs to track stories that only match their interests by selecting a specific path (topic thread) and ignoring the others, which is expected to make the video browsing and the understanding of the topics-of-interest efficient.

4. Analysis of the Topic Thread Structure

4.1 Experiment

In order to analyze the effect of the proposed topic threading



(a) Topic thread structure

(b) The topic thread structure in (a) represented by topic clusters

Fig. 5 A topic thread structure originating from the first story on May 1, 2003 with a search period of $d = 100$ and a relation threshold of $\theta_{trk} = \theta_{cls} = 0.40$.

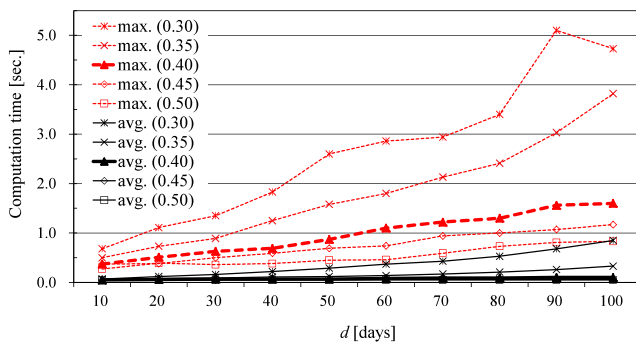
Table 1 Actual contents of the topic clusters in Fig. 5 based on manual analysis.

Cluster #	Contents
$C_T(1)$	SARS found and spreads in Beijing.
$C_T(2)$	SARS spreads all over China.
$C_T(3)$	WHO sends a mission to Beijing.
$C_T(4)$	SARS starts calming down in China, but spreads in Taiwan.
$C_T(5)$	Chinese government worries about the spread of SARS in the countryside.
$C_T(6)$	Chinese government keeps an eye on the spread of SARS in the countryside.
$C_T(7)$	Taiwanese doctor found infected to SARS after a trip to Japan.
$C_T(8)$	Search for SARS infection in Japan.
$C_T(9)$	Anti-SARS conference held in Beijing
$C_T(10)$	SARS calms down in China, but found in Toronto.
$C_T(11)$	SARS calms down in Taiwan, and WHO declares the end of the epidemic.

method, we applied it to actual broadcast news video. The stories that compose each topic thread structure were limited to a search period of $d = 100$ days from the initial story, to

Table 2 Data set and parameters used in the topic thread structure analysis.

News show	NHK News7 (Daily evening news in Japanese. 20 to 30 minutes long)
Period of initial stories	March 1 to June 30, 2002
Search period (d)	100 days; Approximately 171,600 secs. or 47.7 hours of video
Relation threshold (θ_{trk})	0.40

**Fig. 6** Maximum and average computation time of the topic threading process by changing the search period (d) and the relation threshold (θ_{trk} ; the numbers in the parentheses).

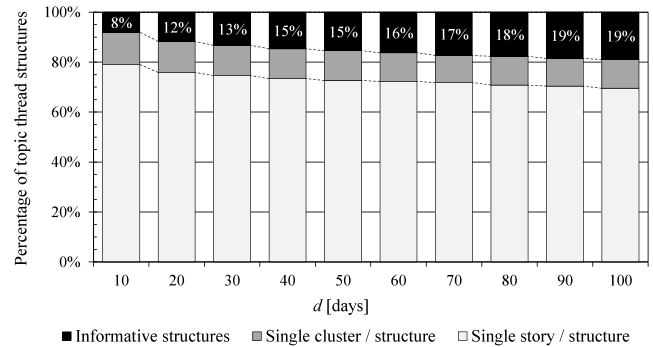
limit the computation time[†]. Table 2 shows the size of the data set and the parameters for the process. The data set is part of the NII TV Broadcast Video Research Corpus: TV-RECS [14], which is composed of more than 1,800 hours of news video archived in the last ten years.

After rejecting stories with only one sentence from the results of the story segmentation, 1,431 initial story (S_O) candidates remained. Next, out of the topic thread structures (T_r) originating from all the initial story candidates, we selected 436 of them that consist of more than two stories. As shown in Fig. 6, under the conditions in Table 2, it took at most a few seconds to obtain a topic thread structure on a Sun Blade 1000 work station, depending on its complexity. The following analyses are performed on these topic thread structures.

Since the computation time roughly reflects the complexity of the topic thread structures, we can see from Fig. 6 that the complexity starts to increase exponentially at $d = 100$ in between $0.40 \leq \theta_{\text{trk}} \leq 0.45$. In general, this indicates that the structure is becoming too complex, and usually, noisy. On the other hand, as in the case of higher values of θ_{trk} , the fact that the complexity does not increase significantly in proportion to d indicates that the topic thread structures are simple and chopped-up into short periods. So we considered that the combination of $d = 100$ and $\theta_{\text{trk}} = 0.40$ is a reasonable setup for the experiment to obtain both informative and less noisy topic thread structures.

4.2 Analysis on the Contents of the Structures

Among the obtained topic thread structures, there were those that revealed informative and comprehensible rela-

**Fig. 7** Percentage of informative and two kinds of not informative topic thread structures by changing the search period (d).

tions between news topics in the real world, such as those exemplified in Figs. 2 and 5.

On the other hand, there were also not informative or incomprehensible structures. Below, we will analyze and discuss such structures.

4.2.1 Not Informative Topic Thread Structures

First of all, in the following experiment, we only considered topic thread structures that were composed of more than one story; structures with only one story were discarded since they do not represent any relation.

However, even among the topic thread structures that were composed of more than one stories, there were structures composed of a single topic cluster. We considered these as “not informative” structures, since they are similar to the results obtained by traditional topic clustering methods that do not represent the transition of topics along time.

Out of the 1,431 topic thread structure candidates, 995 (69.53%) were discarded since they were composed of a single story, 165 (11.53%) were considered not informative since they were composed of a single topic cluster, and as a result, 271 (18.94%) remained as informative ones. The surprisingly small number of informative topic thread structures may be explained by the nature of the daily news show that it tends to take up most topics only when they are strongly focused during a short period of time.

Figure 7 shows the ratio of informative and the two kinds of not informative topic thread structures by changing the search period d . We can see that the ratio of informative structures saturates as d increases. This trend can be explained by the fact that the number of potentially long topic thread structures chopped-up by limiting the search period decreases as d increases. At the same time, it is also explained by the fact that most topics are not discussed for a very long period of time in a news show.

[†]If longer computation time is acceptable depending on the application, we can set longer search periods for d , which will prevent chopping-up potentially long topic thread structures.

4.2.2 Incomprehensible Topic Thread Structures

The incomprehensible structures tend to be caused by mistakenly chaining actually unrelated stories. This occurred firstly when the story segmentation failed, and multiple topics were mixed in a single story. Note that usually, the incomprehensible part is limited to a portion of the topic thread structure, so it does not mean that the entire structure is incomprehensible. Thus, we analyze the incomprehensible structures per topic threads, rather than the entire topic thread structure. Here, a path that connects the initial node S_O and a leaf node S_D is counted as one topic thread. For example, in the structure shown in Fig. 5, there are at least six[†] topic threads connecting S_O and $S_D(j)$ where $j = 1, \dots, 6$.

We selected 76 relatively complex topic threads (those composed of more than eight topic clusters) out of the 1,783 topic threads that appeared in the 436 topic thread structures. The 76 topic threads were composed of 644 stories, of which, 88 were unique. As a result of manual evaluation of the stories, three stories (3%) that appeared in nine topic threads (12%) contained totally unrelated topics due to story segmentation failures. These were considered as incomprehensible topic threads.

Secondly, they occurred when the value of the relation threshold θ_{trk} was inappropriate for the topic; The relation tends to be high for topics on periodic events such as typical crimes or weather phenomena. The appropriateness of the relations between stories is highly dependent on the interest and the motivation of the topic tracking task; for example, tracking how similar events occur along the time (e.g. local crimes in general), or tracking how a specific topic developed along the time (e.g. a specific homicide incident). Thus, we cannot perform a general evaluation and comment on this kind of structure.

Even though there are incomprehensible topic thread structures due to the above reasons, we consider that as long as the structures are used in combination with the mediaWalker interface that will be introduced in Sect. 5, it is easy for a user to notice an incomprehensible structure, and to easily track back to the comprehensible part of the structure.

Since it is also nearly impossible to generate a ground truth structure manually, and also to measure/control the background knowledge of a user, it is difficult to further evaluate the general quality of the obtained topic thread structures. Thus, we will focus on the quantification of the efficiency of browsing the contents along the structure in the next section.

4.3 Analyses on the Size of the Structures

In order to quantify the efficiency of browsing a news video archive based on the topic thread structures, we analyzed them by the number of stories or clusters that composed them, and also by the video length that a user needs to watch to grasp the contents. Since the worst assumption is that a user needs to watch all the video footages in the archive dur-

ing a certain period, we will observe the ratio of the video length in the structures against that of the archive.

4.3.1 Size of Topic Thread Structures

First we analyzed the size of the 436 topic thread structures. Figure 8 shows the histogram of the number of stories and clusters that composed a topic thread structure. From the result, we confirmed the necessity of automatically analyzing the structures, since many of them were composed of many stories and multiple topics, which could be difficult to be analyzed manually.

4.3.2 Size of Topic Threads

Next we analyzed the size of the 1,783 topic threads in the 436 topic thread structures.

Figure 9 shows the histogram of the number of stories and topic clusters that composed a topic thread. As mentioned in Sect. 4.2, quite a number of topic threads were composed of only a few stories, but there were a considerable number of longer ones that indicated the usage of the proposed method should be effective. Meanwhile, the maximum number of topic clusters was nine even for very

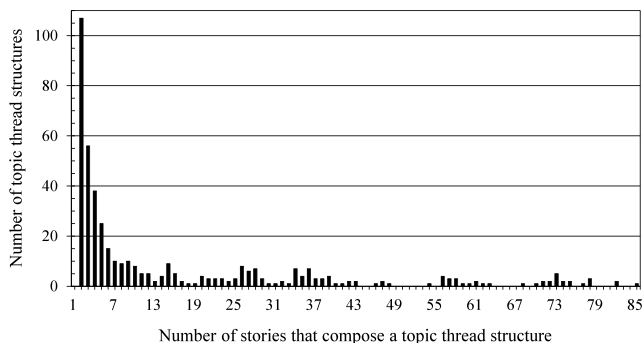


Fig. 8 Histogram of the number of stories that compose a topic thread structure. Note that a structure composed of a single story was discarded.

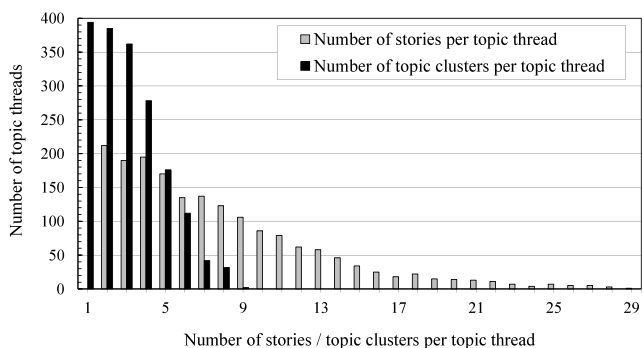


Fig. 9 Histogram of the number of stories/topic clusters that compose a topic thread. Note that topic threads that consist of only one story was filtered out.

[†]Precisely speaking, there are actually more due to multiple paths that connect them.

long topic threads, which leads to the discussion in the next part of this section that using them could make the browsing drastically effective.

Figure 10 shows the histogram of the duration of the topic threads. We assume that the frequency for 100 days was relatively high since there were potentially longer topic threads forced to terminate at d days.

4.3.3 Efficiency of Watching Video Footages along the Topic Thread Structure

In this paper, we assume that a user searches and watches video footages to understand the development of certain topics-of-interest. In order to assist such searching, we provide an interface that allows a user to watch video footages along a topic thread structure, which will be introduced later.

Table 3 shows the effect of watching video footages on the user’s topic-of-interest along a topic thread structure as the statistics of time required and reduced to watch them. As a worst-case baseline, we considered that a user needs to watch all the video footages during a search period d . In the Table, “Topic thread structure” shows the statistics on the total duration of all the video footages in a topic thread structure. It shows that a topic thread structure covers on average 2.6%, and in the worst case 22%, of all the video footages during the period. This case is considered as close to the traditional story-based video retrieval methods that required a user to watch all the video footages that matched a query, though the proposed topic thread structure covers more stories than them since it also includes stories which are tracked from the stories that matched the initial query.

In actual use, we consider that a user would follow a certain topic thread through the browsing interface proposed later in Sect. 5. In order to measure the effect of watching

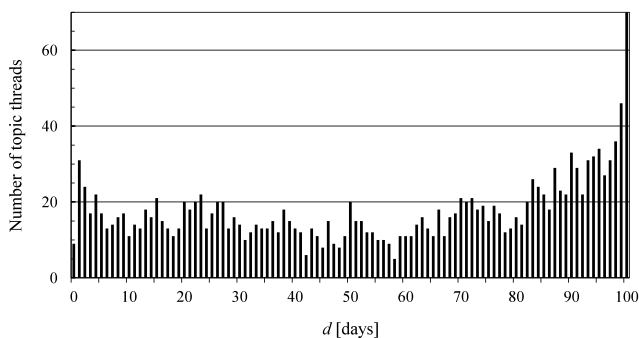


Fig. 10 Histogram of the duration of topic threads.

Table 3 Statistics on the total video length [secs.] of stories that compose the topic structures. The number in parentheses indicates the ratio of the total video length of the stories that compose a topic structure to that of the corresponding search period $d = 100$ days; Approximately 47.7 hours.

	Minimum	Average	Maximum
Topic thread structure	5 (0.0029%)	4,380 (2.6%)	38,125 (22%)
Topic thread	5 (0.0029%)	2,770 (1.6%)	19,038 (11%)
Topic cluster	5 (0.0029%)	1,280 (0.75%)	6,741 (3.9%)

the video footages along the topic threads, in Table 3, “Topic thread” shows the statistics on the total time of all the video footages along a topic thread, and “Topic cluster” shows that of only the video footages of the stories which were the cluster centers (C_0 ; The first story in a topic cluster). The results show that browsing along a topic thread reduces the time needed to watch video footages related to the user’s topics-of-interest by 1% on average and 11% in the worst case. Browsing by topic clusters along a topic thread, which could be considered as a digest view of the topic thread, further reduces it to 0.85% on average and 7.1% in the worst case. We consider the reductions should be extremely effective for practical use.

5. mediaWalker II: A News Video Browsing and Editing Interface

5.1 Overview of the Interface

We have previously implemented an interface for browsing and editing news video based on the topic thread structure named “mediaWalker”[†] [2], [15]. Here, we introduce an extended version of the interface that supports linking to external resources; “mediaWalker II”. The interface is composed of the following sub-interfaces:

- Initial story search/listing interface
- Initial story selection interface
- Topic thread browsing interface

Details of the functions of each sub-interface are described step-by-step in the following sections.

5.2 Initial Story Search/Listing Interface

To start with, a user needs to select an initial story. As shown in Fig. 11, three options are provided through the *Initial story search/listing* interface; 1) Select a pre-defined story set, 2) Search for a story by a combination of a query

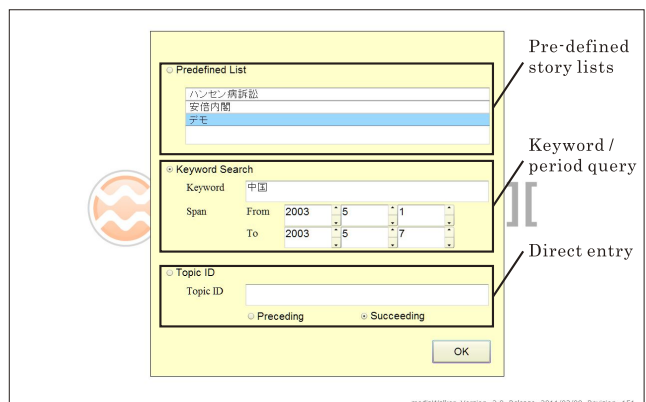


Fig. 11 The mediaWalker II interface: Initial story search/listing sub-interface.

[†]The interface is demonstrated at the following URI: <http://www.murase.m.is.nagoya-u.ac.jp/~ide/res/mediaWalker/>.

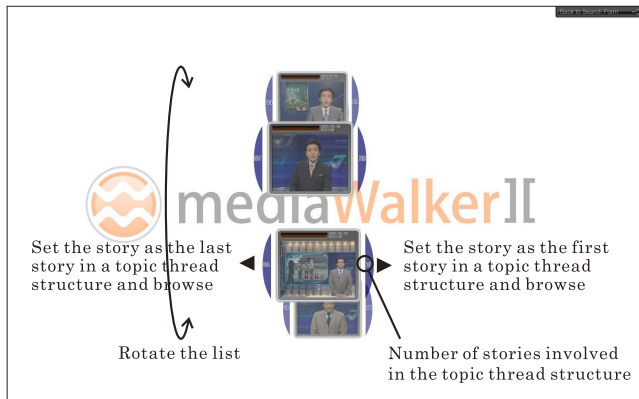


Fig. 12 The mediaWalker II interface: Initial story selection sub-interface.

term and a period of time, or 3) Directly enter a story ID. Since the query term field can be left blank, listing all stories during a certain period is also possible in method 2). Stories from the selected story set, the result of the query, or the entered ID are listed in the subsequent *Initial story selection* interface.

5.3 Initial Story Selection Interface

As shown in Fig. 12, in the *Initial story search/selection* interface, thumbnail images of stories are listed. The stories listed here are those that matched the criterion specified in the *Initial story listing* interface. The list of thumbnail images could be rotated, and also played when clicked. Once finding an interesting story, the user sets it as the initial story by pressing either the “Past” or the “Future” buttons next to the thumbnail; the former will show the topic thread structure originating from the story towards the past, and the latter towards the future in the subsequent *Topic thread browsing* interface.

The size of the topic thread structure is shown for reference on the buttons as the number of stories composing the structure.

5.4 Topic Thread Browsing Interface

5.4.1 Topic Tracking Functions

As shown in Fig. 13, the *Topic thread browsing* interface visualizes the topic thread structure originating from the initial story, where each node (story) is represented by a thumbnail image. Topic clusters are shown as hatched regions in the background. A user will track down a topic thread structure by clicking the thumbnail images to watch the corresponding video footage story after story. This function allows a user to efficiently explore a news video archive based on the context, making use of the characteristics of the topic structures described in Sect. 3. An example of a user tracking along a topic thread is shown in Fig. 14.

In addition, each thumbnail image is accompanied by

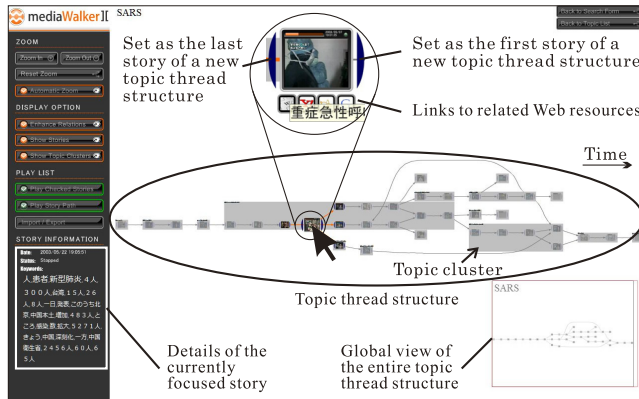


Fig. 13 The mediaWalker II interface: Topic thread browsing sub-interface.

a pair of parentheses, which allows the user to set the story as a new initial story and show its topic thread structure to the past (left parenthesis) or to the future (right parenthesis). This function allows the user to freely explore the entire archive back and forth along the topic thread structures.

5.4.2 External Link Function

Furthermore, we have implemented a function that allows users to access external contents by clicking icons that may appear below the thumbnail image, according to availability. Currently, the links are limited to Web contents, which are established by first linking to Wikipedia articles according to the method proposed by Okuoka et al. [16], and next to other Web contents by issuing queries based on the title of the linked Wikipedia articles or frequent terms that appear in each story.

5.4.3 Video Editing (“Story Telling”) Function

The *Topic thread browsing* interface is equipped with another function that allows a user to select stories to “tell stories” by rearranging the video footages in the archive.

As shown in Fig. 15, besides simply selecting each story, this function allows a user to select stories along a topic thread between two or more specified stories. The interface outputs the selected stories as a list of story/video IDs, which could be exported as an input to a video summarization system, or simply as a play-list for a video editing software.

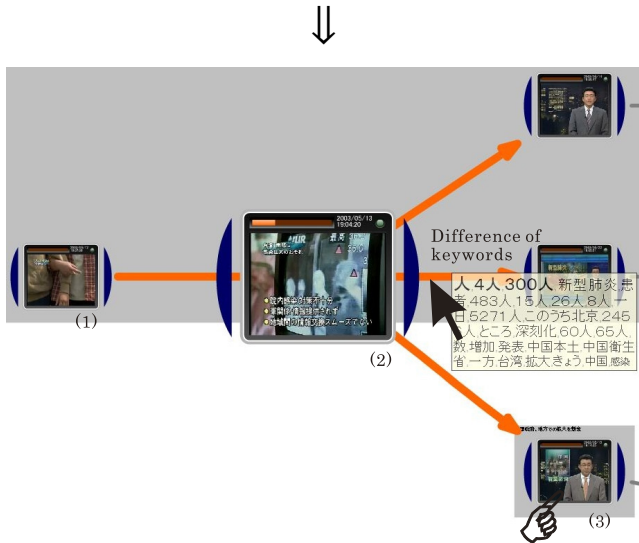
6. Related Works

6.1 Works on the Analysis of Chronological Semantic Structures in News

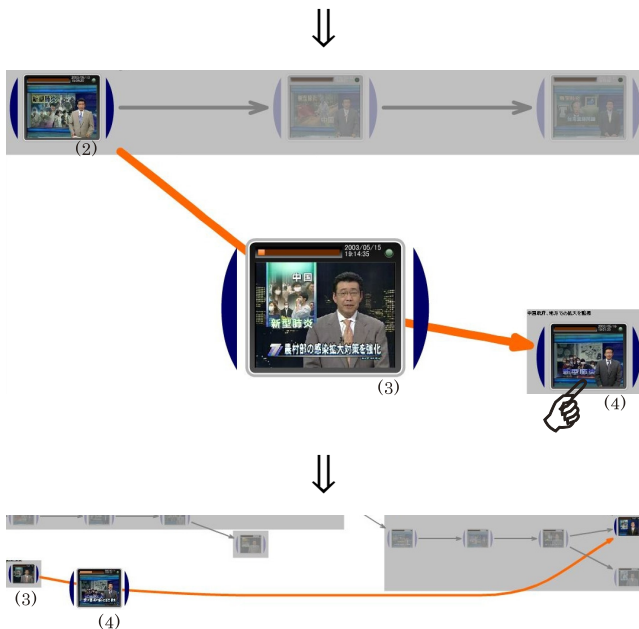
The simplest approach to structure news stories is to first cluster related stories into a topic and then chain them linearly in chronological order, as in the “Topic detection task” defined in the TDT workshop series [6]. A majority of the



(a) When a thumbnail image that represents a story is clicked, it is enlarged and played. The enlargement is automatically adjusted so that the preceding and the succeeding stories should remain in the screen.



(b) When the pointer overlaps an edge, an overlaid window appears and shows the difference of keywords between the stories on both sides of the edge. This is helpful especially when selecting a topic thread at a diverging story.



(c) The user will keep on selecting and watching video footages of stories along a topic thread structure in this manner. This helps the understanding of the topics present in the topic thread structure.

Fig. 14 Tracking down story by story in the *Topic thread browsing* interface, along the topic thread structure shown in Fig. 5.

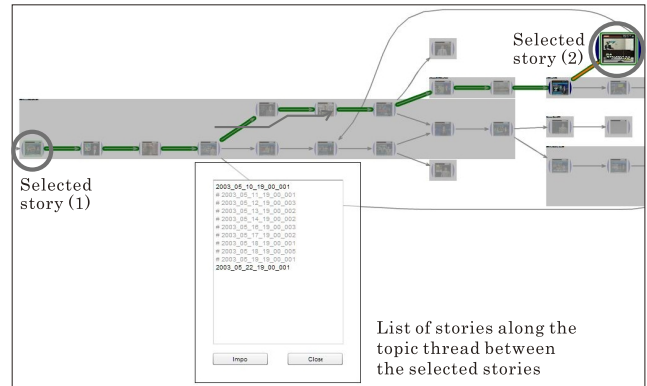


Fig. 15 The mediaWalker II interface: Video editing and exporting function.

existing works on news story structuring is based on this approach, such as Duygulul et al.’s work on news video tracking based on the detection of similar images and logos [17]. However, when the size of the data set explodes, this approach is not appropriate in the following two senses. First, since the structure tends to extend in a long chain of stories, watching video footages along the structure is extremely time consuming. Second, a linear structure could not express the simultaneous flow of multiple topic threads within a topic-of-interest.

In the final TDT workshop (TDT2004), the “Hierarchical topic detection task” [18] was introduced. It aimed at structuring a directed acyclic graph of stories instead of clusters, but it was more of a way to assign multiple topics to a story and describe relations between topics. Meanwhile, Wu et al. proposed a method that structures clustered news stories into a binary tree based on the chronological order and the local change in a topic [19]. This method partially solves the above-mentioned problems, but it still could not represent the simultaneous flow of multiple topic threads.

The method proposed in this paper structures stories into a graph so that it could represent such simultaneous flows. As related works that share the same approach, there are topic threading methods for news paper articles [20]–[22]. However, these methods do not consider the independence between topic threads at a diverging node, which results in the existence of redundant edges between nodes, where the proposed method eliminates such edges, for the sake of efficient tracking.

Another difference of the proposed method with the majority of the existing works, except for Nallapati et al.’s work [20], is that it does not cluster a set of stories that forms a topic beforehand. It instead forms the topic thread structure by chaining locally related stories under certain rules and later extracts locally related clusters from the structure, considering the nature of news contents that a topic may gradually develop into, diverge to, or merge with other topics. We consider that such a feature is essential to find unobvious relations between topics such as those shown in Fig. 2.

6.2 Works on the Visualization of News Structure

A typical method to visualize a video structure on a screen is the so-called “storyboard,” which lists up an arbitrary number of representative frames from the video as thumbnail images, for example in Carnegie Mellon University’s original Informedia News-on-Demand system [3], and Dublin City University’s Físchlár News interface [23]. The interface proposed in this paper could be considered as a two-dimensional storyboard represented by a directed graph.

As for the works in visualizing news structure, most of them focus on the so-called 5W1H (When, Where, Who, What, Why, and How) attributes, and especially the first three of them (3Ws). As part of the Carnegie Mellon University’s Informedia project, Christel et al. proposed a news video browsing interface that visualizes news stories based on the combinations of the 3Ws focusing especially on the “Where” attribute [24], [25]. We focused on the cooccurrence of the “Who” attribute, and proposed a news video browsing interface by exploring the social network in news contents [26].

Meanwhile, the proposed interface makes use of the news contents as a whole except for the “When” attribute. It only considers the 3Ws by giving certain weights to them when evaluating the relations between stories. As a related work, Rautiainen et al. proposed a cluster-temporal browsing method [27] which visualizes the cluster of stories with similar image features. On the other hand, de Rooij et al. proposed an interface that maps semantic and chronological relations between news video footages on a sphere or a flat plane [28].

As for supporting authoring video contents in an archive, Casares et al. proposed the “Silver” interface [29], which allows the authoring of a single video, but not the re-compilation of multiple video footages as in the proposed interface.

7. Conclusion

In this paper, we proposed a topic threading method for a large amount of news video, and also introduced an interface that allows the searching, browsing and tracking of news video based on the topic thread structures. Analysis on the contents revealed interesting and comprehensible relations between actual topics in the real world, while analyses on the size of the topic thread structures quantified the efficiency of tracking the development of topics-of-interest and also the selection of video footages for a further authoring purpose.

Although the proposed method was developed based on a Japanese news video archive, the topic threading algorithm and the mediaWalker interface are language independent, since they are purely based on the relation between pairs of stories. Optimal parameters may differ according to each language, but we believe that the proposed approach could be applied to news video archives in other languages.

As future works, we have started working on linking the structures with news shows in other languages [30], other channels, and shared video footages, images, articles on the web, together with the development of a method to “tell stories” by making use of the topic thread structures.

Acknowledgments

Parts of the work presented in this paper were supported by the Grants-in-Aid for Scientific Research (15700116, 18049035, 18700080, 19024036, 21013022, 21700103, 23300036), Kayamori Information Science Foundation (K17KenX 202), Hosono Bunka Foundation (211033), and Mitsubishi Research Institute, Inc. The work is a fruit of a joint research project between Nagoya University, NetCOMPASS Ltd., Gifu Shotoku Gakuen University, and the National Institute of Informatics, Japan.

References

- [1] I. Ide, T. Kinoshita, T. Takahashi, H. Mo, N. Katayama, S. Satoh, and H. Murase, “Exploiting the chronological semantic structure in a large-scale broadcast news video archive for its efficient exploration,” Proc. APSIPA Annual Summit and Conf. 2010, pp.996–1005, Singapore, Dec. 2010.
- [2] I. Ide, T. Kinoshita, T. Takahashi, H. Mo, N. Katayama, S. Satoh, and H. Murase, “Proposal of an information compilation method for massive news video data based on their time-series semantic structure (in Japanese),” Trans. JSAI, vol.23, no.5, pp.282–292, Sept. 2008.
- [3] A.G. Hauptmann, M.J. Witbrock, and M.G. Christel, “News-on-Demand: An application of informedia technology,” D-lib Mag., no.3, <http://www.dlib.org/dlib/september95/nod/09hauptmann1.html>, Sept. 1995.
- [4] S. Bocconi, F. Nack, and L. Hardman, “Automatic generation of matter-of-opinion video documentaries,” J. Web Semantics, vol.6, no.2, pp.139–150, April 2008.
- [5] P. Shrestha, P.H. de With, H. Weda, M. Barbieri, and E.H. Aarts, “Automatic mashup generation from multiple-camera concert recordings,” Proc. Eighteenth ACM Multimedia Conf., pp.541–550, Firenze, Toscana, Italy, Oct. 2010.
- [6] J.G. Fiscus and G.R. Doddington, “Topic detection and tracking evaluation overview,” in Topic Detection and Tracking: Event-based Information Organization, ed. J. Allan, pp.17–31, Kluwer Academic Publishers, Norwell, MA, USA, 2002.
- [7] A.F. Smeaton, P. Over, and W. Kraaij, “Evaluation campaigns and TRECVID,” Proc. Eighth ACM International Workshop on Multimedia Information Retrieval, pp.321–330, Santa Barbara, CA, USA, Oct. 2006.
- [8] T.S. Chua, S.F. Chang, L. Chaisorn, and W. Hsu, “Story boundary detection in large broadcast news video archives — Techniques, experiences, and trends,” Proc. Twelfth ACM International Multimedia Conf., pp.656–659, New York, NY, USA, Oct. 2004.
- [9] A. Merlino, D. Morey, and M. Maybury, “Broadcast news navigation using story segmentation,” Proc. Fifth ACM Multimedia Conf., pp.381–391, Seattle, WA, USA, Nov. 1997.
- [10] A.G. Hauptmann and M.J. Witbrock, “Story segmentation and detection of commercials in broadcast news video,” Proc. IEEE International Forum on Research and Technology Advances in Digital Libraries 1998, pp.168–179, Santa Barbara, CA, USA, April 1998.
- [11] I. Ide, H. Mo, N. Katayama, and S. Satoh, “Topic-based inter-video structuring of a large-scale news video corpus,” Proc. 2003 IEEE International Conf. on Multimedia and Expo, pp.305–308, Baltimore, MD, USA, July 2003.

- [12] I. Ide, H. Mo, N. Katayama, and S. Satoh, "Topic-based structuring of a very large-scale news video corpus," Tech. Rep. SS-03-08-11, AAAI Press, March 2003.
- [13] I. Ide, R. Hamada, S. Sakai, and H. Tanaka, "Semantic analysis of television news captions referring to suffixes," Proc. Fourth International Workshop on Information Retrieval with Asian Languages, pp.37–42, Taipei, Taiwan, Nov. 1999.
- [14] N. Katayama, H. Mo, I. Ide, and S. Satoh, "Mining large-scale broadcast video archives towards inter-video structuring," Advances in Multimedia Information Processing — PCM2004 Fifth Pacific Rim Conference on Multimedia, Tokyo, Japan, November/December 2004, Proc. Part II —, ed. K. Aizawa, Y. Nakamura, and S. Satoh, Lect. Notes Comput. Sci., vol.3332, pp.489–496, Tokyo, Japan, Springer-Verlag, Berlin, Heidelberg, Dec. 2004.
- [15] I. Ide, T. Kinoshita, T. Takahashi, S. Satoh, and H. Murase, "mediaWalker: A video archive explorer based on time-series semantic structure," Proc. Fifteenth ACM International Multimedia Conf., pp.162–163, Augsburg, Bavaria, Germany, Sept. 2007.
- [16] T. Okuoka, T. Takahashi, D. Deguchi, I. Ide, and H. Murase, "Labeling news topic threads with Wikipedia entries," Proc. Eleventh IEEE International Symposium on Multimedia, pp.501–504, San Diego, CA, USA, Dec. 2009.
- [17] P. Duygulu, J.Y. Pan, and D.A. Forsyth, "Towards auto-documentary: Tracking the evolution of news stories," Proc. Twelfth ACM International Conf. on Multimedia, pp.820–827, New York, NY, USA, Oct. 2004.
- [18] A. Feng and J. Allan, "Hierarchical topic detection in TDT-2004," Tech. Rep. IR-389, Univ. of Massachusetts, Center for Information Retrieval, April 2005.
- [19] X. Wu, C.W. Ngo, and Q. Li, "Threading and autodocumenting news videos," IEEE Signal Process. Mag., vol.23, no.2, pp.59–68, March 2006.
- [20] R. Nallapati, A. Feng, F. Peng, and J. Allan, "Event threading within news topics," Proc. Thirteenth ACM Conf. on Information and Knowledge Management, pp.446–453, Washington, DC, USA, Nov. 2004.
- [21] N. Uramoto and K. Takeda, "A method for relating multiple newspaper articles by using graphs, and its application to Webcasting," Proc. 36th Annual Meeting of the Association for Computational Linguistics and the 17th International Conf. on Computational Linguistics, pp.1307–1313, Montréal, QC, Canada, Aug. 1998.
- [22] C.C. Yang, X. Shi, and C.P. Wei, "Tracing the event evolution of terror attacks from on-line news," Intelligent and Security Informatics — IEEE International Conference on Intelligence and Security Informatics, ISI 2006, San Diego, CA, USA, May, 2006, Proc. —, ed. S. Mehrotra, D.D. Zeng, H. Chen, B. Thuraisingham, and F.Y. Wang, Lect. Notes Comput. Sci., vol.3975, pp.343–354, San Diego, CA, USA, Springer-Verlag, Berlin, Heidelberg, May 2006.
- [23] H. Lee, A.F. Smeaton, N.E. O'Connor, and B. Smith, "User evaluation of Físchlár-News: An automatic broadcast news delivery system," ACM Trans. Information Systems, vol.24, no.2, pp.145–189, April 2006.
- [24] M.G. Christel, A.M. Olligschlaeger, and C. Huang, "Interactive maps for a digital video library," IEEE Multimedia, vol.7, no.1, pp.60–67, Jan.–March 2000.
- [25] M.G. Christel, A.G. Hauptmann, H.D. Wactler, and T.D. Ng, "Collages as dynamic summaries for news video," Proc. Tenth ACM International Conf. on Multimedia, pp.561–569, Juan-les-Pins, Alps-Maritimes, France, Dec. 2002.
- [26] I. Ide, T. Kinoshita, H. Mo, N. Katayama, and S. Satoh, "trackThem: Exploring a large-scale news video archive by tracking human relations," Information Retrieval Technology, Second Asia Information Retrieval Symposium, AIRS2005, Jeju Island, Korea, Oct. 2005, Proc., ed. G.G. Lee, A. Yamada, H. Meng, and S.H. Myaeng, Lect. Notes Comput. Sci., vol.3689, pp.510–515, Jeju, Korea, Springer-Verlag, Berlin, Heidelberg, Oct. 2005.
- [27] M. Rautiainen, T. Ojala, and T. Seppänen, "Cluster-temporal brows-

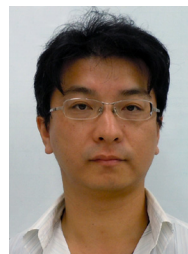
ing of large news video databases," Proc. 2004 IEEE International Conf. on Multimedia and Expo, pp.751–754, Taipei, Taiwan, June 2004.

- [28] O. de Rooij and M. Worring, "Browsing video along multiple threads," IEEE Trans. Multimed., vol.12, no.2, pp.121–130, Feb. 2010.
- [29] J. Casares, B.A. Myers, A.C. Long, R. Bhatnagar, S.M. Stevens, L. Dabbish, and A. Corbett, "Simplifying video editing with intelligent interaction," Adjunct Proc. CHI2002: Human Factors in Computing Systems, pp.672–673, Minneapolis, MN, USA, April 2002.
- [30] A. Ogawa, T. Takahashi, I. Ide, and H. Murase, "Cross-lingual retrieval of identical news events using image information," Advances in Multimedia Modeling — Fourteenth International Multimedia Modeling Conference, MMM2008, Kyoto, Japan, Jan. 2008, Proc., ed. S. Satoh, F. Nack, and M. Etoh, Lect. Notes Comput. Sci., vol.4903, pp.287–296, Kyoto, Japan, Springer-Verlag, Berlin, Heidelberg, Jan. 2008.



Ichiro Ide received his B.Eng., M.Eng., and Ph.D. from The University of Tokyo in 1994, 1996, and 2000, respectively. He became an Assistant Professor at the National Institute of Informatics, Japan in 2000. Since 2004, he has been an Associate Professor at Nagoya University. He had also been a Visiting Associate Professor at National Institute of Informatics from 2004 to 2010, an Invited Professor at Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA), France in 2005, 2006, and

2007, a Senior Visiting Researcher at Instituut voor Informatica, Universiteit van Amsterdam from 2010 to 2011. He is also a member of IPS Japan, ITE, JSAI, ANLP, IEEE, and ACM.



Tomoyoshi Kinoshita received his B.Eng., M.Eng., and Ph.D. from The University of Tokyo in 1995, 1997, and 2000, respectively. He has been the President of NetCOMPASS, Ltd. since 2000. He is also a member of IPS Japan, JSAI, IEEE Computer Society, and ACM.



Tomokazu Takahashi received his B.Eng., M.Eng., and Ph.D. from Ibaraki University, in 1997, 2000, and 2003, respectively. He became a Post Doctoral Researcher for the 21st Century COE program at Nagoya University from 2003, and a JSPS Post Doctoral Research Fellow from 2005. Since 2008, he has been an Associate Professor at Gifu Shotoku Gakuen University. He is also a member of IEEEJ.



Hiroshi Mo received his B.Eng. and M.Eng. from Musashi Institute of Technology in 1990 and 1992, respectively, Ph.D. from The University of Tokyo in 1997. He became an Assistant Professor in 1998 at National Center for Science Information Systems, Japan. Since 2000, he has been an Assistant Professor at National Institute of Informatics, Japan. He is also a member of IPS Japan, SOFT, and ITE.



Norio Katayama received his B.Eng., M.Eng., and Ph.D. from The University of Tokyo in 1990, 1992, and 1995, respectively. He became an Assistant Professor in 1995 at National Center for Science Information Systems, Japan. Since 2000, he has been an Associate Professor at National Institute of Informatics, Japan. He is also a member of IPS Japan, IEEE, and ACM.



Shin'ichi Satoh received his B.Eng., M.Eng., and Ph.D. from The University of Tokyo in 1987, 1989, and 1992, respectively. He became an Assistant Professor and Associate Professor in 1992 and in 1998, respectively at National Center for Science Information Systems, Japan. In 2000, he became an Associate Professor at National Institute of Informatics, Japan. Since 2004, he is a Professor at National Institute of Informatics, Japan. He had also been a Visiting Researcher at Carnegie Mellon University from 1995 to 1997. He is also a member of IPS Japan, IEEE Computer Society, and ACM.



Hiroshi Murase received his B.Eng. and M.Eng., Ph.D. from Nagoya University in 1978, 1980, and 1987, respectively. He had been a researcher at NTT from 1980 to 2003, and also a Visiting Researcher at Columbia University in 1992. Since 2003, he has been a Professor at Nagoya University. He is also a member of IPS Japan and an IEEE fellow.