# Expansion of Multimodal Summarization for Trend Information – Report on the First and Second Cycles of the MuST Workshop –

Tsuneaki Kato
The University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
kato@boz.c.u-tokyo.ac.jp

Mitsunori Matsushita
NTT Communication Science Laboratories, NTT Corp.
2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan
mat@clsab.kecl.ntt.co.jp

Noriko Kando
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
kando@nii.ac.jp

## Abstract

*A trend, which is a general tendency in the way a situation is changing or developing, is based on temporal statistical data, and can be obtained by synthetically summarizing it. A system that can properly answer questions on interesting trends must help users to understand and review a huge amount of information. MuST, a workshop on multimodal summarization for trend information, was designed to encourage cooperative and competitive studies on multimodal summarization for trend information, and to allow it to construct such an information system. MuST is a pilot workshop of the NTCIR workshop. It has been run based on approximately one year as one cycle, and it finished its second cycle in March 2007. Several researches with broader scope than initially expected have been conducted in the workshop. This paper explains the framework and objective of the MuST workshop, reviews researches conducted there and the results obtained so far, and discusses the role of the workshop for those researches.*

## 1 Introduction

A trend, which is a general tendency in the way a situation is changing or developing (cf. Longman Advanced American Dictionary), is based on temporal statistical data, and can be obtained by synthetically summarizing it, but not by simple enumeration. For understanding and reviewing a huge amount of information, it would be very helpful to be supported by a system that can properly answer questions on interesting trends, such as how the price of gasoline changed during the year, what the situation of the domestic personal computer market has been over the last few years, and how severe the typhoons were last autumn.

In considering such a system for compiling trend information, we notice that its input consists of non-linguistic information, that is, temporal data spanning a given period, and linguistic information consisting of explanations and evaluations of that data. The output consists of both plain and concise text and non-linguistic information, such as information graphics, that are used to help users reach a more effective understanding. These considerations reveal the need for a technology that uniformly summarizes linguistic and non-linguistic information and compiles it into multimedia presentations in which text and visual information works cooperatively. We call this new technology *multimodal summarization*. [1]

From the viewpoint of research communities on natural language processing and text processing, multimodal summarization can be seen as an effort to expand text summarization, which extracts important content from a huge amount of real-world text and presents it in condensed form, in order to include into its targets non-linguistic information such as numerical data and information graphics. Whereas multimedia presentation generation, which has been actively studied for decades, aims to generate multime-

[1] we avoid the term, *multimedia summarization*, because it is commonly used to summarize audio and visual information.

dia presentations from media-independent semantic representations, multimodal summarization does not presume the existence of such a well-formed semantic representation and tackles the massive amount of unstructured and uncoordinated information available in the real world. Although the target of multimodal summarization is not limited to trend information, handling trend information is an excellent starting point since, as mentioned before, it needs to take both linguistic and non-linguistic information into account as its input and also as its output.

Multimodal summarization for trend information needs the creative fusion of researches in several fields. Such fields include multimedia presentation generation, text summarization, information visualization, information extraction, and question answering, among some of which there has been little interaction so far. MuST, a workshop on multimodal summarization for trend information, has been proposed in order to encourage the fusion and progress of researches in those fields and to foster researches on multimodal summarization for trend information [1]. This paper explains the MuST workshop and describes results obtained so far. First, in section 2, its framework and objectives are explained, and the course taken so far is reviewed. Section 3 describes the MuST data set which plays an important role in the workshop framework. Then, section 4 presents several research activities conducted in the MuST workshop. Section 5 discusses the role of the workshop in those researches. Finally, a future direction is suggested and the whole paper is concluded.

## 2 Framework and Course

The MuST workshop was designed to encourage cooperative and competitive studies on multimodal summarization for trend information. One purpose of the workshop is to promote discussion and community conformity, while another is to construct and accumulate resources such as tools and corpora. This will be encouraged by sharing the same research resource and by conducting research concerning a common theme in a broad sense. An objective and quantitative evaluation using a test set is outside the scope of the present workshop, which distinguishes this workshop from other NTCIR tasks.

The major characteristic of the MuST workshop is the use of a shared resource, which motivates participation in the workshop and substitutes for objective evaluation using a common test set in evaluation workshops. The resource that is shared, the MuST data set, includes not only the materials to be processed, but also the intermediate results behaving as a hub, and references for the output design. These intermediate results behaving as a hub mean that it can be the output of some modules and also the input of other modules within a multimodal summarization system. Although multimodal summarization requires several component technologies that are dispersed in many research fields, which makes it difficult to construct a total system, by using this data set the participants can address their own subjects. This is especially important for those studying elemental technologies. Moreover, participants from different communities can discuss their interests with each other using the data set as common ground, and can see how their studies or their modules fit into the whole framework. Of course, researchers having the same interest can use the data set as a material for objective evaluation. Encouraging and fostering researches through such interchanges is the objective of sharing a research resource and of the MuST workshop.

MuST was announced in November 2004. In early 2005, a mailing list of interested persons was set up, a call for participants was started, and we began to deliver parts of the data set to the participants. The workshop runs with a cycle of around one year. In the first cycle, MuST had fifteen participants. A two-day closed workshop was held in March 2006 to wrap up the first cycle. The second cycle gathered eighteen participants and, like the first cycle, a two-day closed workshop was held in March 2007. In both workshops, all participants gave a presentation on their research progress and held intensive discussions. Table 1 shows the participants. A roundtable meeting was held in each cycle besides the final workshop. In addition to the design, construction and maintenance of the data set and organization of those meetings, the roles of the organizers include providing information through the home page and mailing list, and encouraging the accumulation and sharing of resources such as data and tools made by participants.

## 3 The MuST Data Set

A multimodal summarization system for trend information takes a topic as an input from a user who is interested in it, collects related information, compiles it to make trend information, and represents it as a multimedia presentation. The MuST data set was designed and constructed with the objective of becoming a shared research resource for fostering the development of such a system and research on their component technologies.

The MuST data set consists of data used to compile trend information, intermediate results of the compilation process, and examples/references of components of multimedia presentations to be obtained on several topics, each of which corresponds to a users' query or information request. It has a wide range of topics: social and economic domains or activities, such as the oil industry, the personal computer market and car production; groups of incidents such as earthquakes,

**Table 1. Participants in the MuST workshop**

1. A.I. Lab., Department of Information Engineering, Faculty of Engineering, Mie University */**
2. Communication Design Lab., Department of Information Systems and Multimedia Design, School of Engineering, Tokyo Denki University **
3. Computational Intelligence Lab., Department of Computer Science and Intelligence Systems, Graduate School of Engineering, Osaka Prefecture Universiry */**
4. Computational Linguistics Group, National Institute of Information and Communications Technology */**
5. Corporate Research & Development Center, Oki Electric Industry Co., Ltd *
6. Department of Computer Science, Graduate School of Natural Science and Technology, Okayama University */**
7. Department of Mathematics and Information Sciences, Graduate School of Science, Osaka Prefecture Universiry */**
8. Faculty of System Design, Tokyo Metropolitan University */**
9. Fukumoto Lab., Ritsumeikan University **
10. Graduate School of Information Sciences, Hiroshima City University */**
11. Innovative Technology R&D Department, Justsystems Corporation */**
12. Itoh Lab., Department of Information Sciences, Ochanomizu Universiry **
13. Knowledge Data Engineering Laboratory, Toyohashi University of Technology *
14. Kobayashi Lab., Ochanomizu University */**
15. Management Information System Group, Graduate School of Engineering, Osaka Prefecture University **
16. Media Technology Research Center, Dai Nippon Printing Co., Ltd *
17. Mori Lab., Graduate School of Environment and Information Sciences,Yokohama National University */**
18. Murata Lab., Department of Computer Science, Faculty of Engineering, Tokyo Institute of Technology **
19. NTT Communication Science Laboratories & The University of Tokyo */**
20. Saito Lab., Department of Information and Computer Science, Keio University */**
21. Service Platforms Research Laboratories, NEC **

* First cycle, ** Second cycle, */** First and second cycles

typhoons, and car accidents; and organizations such as Sony Corp. and the Obuchi Cabinet. Those trends can be compiled from linguistic descriptions of statistics and/or reports on events. For example, trends in the personal computer industry include statistics on shipment volume, shipment value, and market share of major makers. Typhoon trends consist of a review of typhoon-related events, such as their development and landfalls, and damage statistics. In the MuST data set, these descriptions of statistics and events are considered units of information. In addition to those primitives, situation descriptions and evaluations found in, say, newspaper articles are materials for compiling trend information.

The current MuST data set makes much use of linguistic information and uses as its document set the Japanese Mainichi newspapers from 1998 and 1999. We selected 27 topics, and selected about three related statistics and event types for each topic, 77 in total. For each topic, we collected and constructed the following data:

- A list of newspaper articles (582 articles in total) concerning the topic
- Annotations of those articles
- Examples of textual summaries on the selected statistics and events
- Tables or charts on those statistics

*List of articles* provides articles obtained through the steps for collecting information. *Textual summaries* are examples of the output of the summarization of each statistic. *Tables and charts* provide data collected and/or examples of parts of summarization results.

*Annotated articles* is the most important data in the MuST data set. Examples in which the annotation schema was applied to English texts are shown in Figure 1, instead of examples of real data, which is in Japanese.

First, sentences mentioning selected statistics or events are annotated as `unit` elements, each of which has a `stat`/`event` attribute that shows which statistic or event it is related to. Most annotations are conducted on sentences selected as `unit` elements. To make them understandable after extraction from the surrounding context, ellipses (zero anaphora) and omitted phrases are supplied and marked as `ins` elements. `Ins` elements have an attribute that indicates what elements such as the name of the statistic and the dates of the report are supplied. Some clauses in `unit` elements are marked as `del` elements when their parts do not relate directly to the statistics or events con-

<unit stat="nation wide average of pomp price of gasoline"><del type="src">Based on the July 6th report announced by the Oil Information Center,</del> <name part="head">the price of gasoline (one liter, regular)</name>, based on the research conducted <date gra="week" abs="19990617">this week</date>, reached a <name part="foot">national average</name> of <val>92 yen</val>, <rel type="diff">1 yen</rel> higher than <date gra="week" abs="19990610">last week</date>'s <name part="head">average price</name></unit>.

<unit stat="Dubai oil price"><name>The oil price (Dubai Oil)</name> has kept dropping since its <rel type="ord">peak</rel> <date gra="month" abs="199710">last October</date>, of <val>around $20</val> <name part="foot">per barrel</name>, and fell to <val>$12.50</val>. in <date gra="ten-days" abs="19980121">late January</date></unit>. <unit stat="Dubai oil price">After <date gra="ten-days" abs="19980121">that</date>, ø<ins type="name">Dubai oil price</ins> rose temporarily, <del type="rsn">because of the tension of the Iraq situation,</del> but has been struggling recently at <val>around $10</val><del type="other">, being "on the oversupplied side"</del></unit>.

<unit event="typhoon landfall">Medium-strength <par>typhoon No. 10</par> struck <par>Makuraszaki-shi, Kagoshima</par>, at <date gra="hour" abs="19980917">about 4:30 pm on the 17th</date>, and will strike in <par>the vicinity of Shukumo-shi, Kochi</par> <date gra="hour" abs="19980917">the same night</date></unit>.

<unit stat="domestic shipment volume"><name>The domestic shipment volume </name>for <date gra="half-year" abs="199804">the first half of the year</date> was <val>4,391,000</val>, which is <rel type="prop">34%</rel> higher than for <date gra="half-year" abs="199704">the same period last year</date> and marked <rel type="ord">the highest level</rel> for <name part="foot">the half-year range</foot></unit>.

**Figure 1. Examples of Annotations on English Text**

cerned. They include the sources of information and the impact of the value change or event, which are categorized and described as an attribute of the `del` element.

From the text of a `unit` element, phrases mentioning the name of the statistic (`name` element), the value of the statistic (`val` element), the relative values, which relate to the statistic but are not the value itself (`rel` element), dates (`date` element), and other parameters (`par` element) are determined and marked as the respective element (shown in parentheses). `Name` elements have an attribute that shows which statistic that element denotes, since there is sometimes more than one statistic mentioned per `unit` element. In addition, since the name of a statistic can be expressed in a variety of ways and can be scattered in multiple non-consecutive phrases in the sentence, some `name` elements consist of just part of a statistic name. To cope with this situation, `name` elements have an attribute that indicates whether they are complete names, main portions of a name, or complementary portions. `Val` and `rel` elements are annotated to clarify their relationship to a particular statistic; `rel` elements are subcategorized into differences, proportions, and orders, which categorizations are indicated as attributes. `Date` elements have attributes that indicate their granularity and absolute dates when using relative expressions. For example, "Yesterday" is annotated with `gra = "day", abs = "19990203"`. `Par` elements correspond to arguments of the statistic or event, such as companies for market share and the name of a typhoon and a place for typhoon landfalls. Dates are handled separately because of their special position, though they could be a kind of argument. Anaphoric expressions are marked as such, and constitute `ref` elements with an attribute that shows their antecedent. Again, dates are special in this treatment as well.

From the specification above, it can be understood that the annotation of the MuST data set represents the intermediate results of semantic and pragmatic analysis specialized for statistical and/or event information. In the flow of summarization in general, extraction and analysis of important sentences is followed by rephrasing and sentence generation to eliminate redundancy and maintain consistency. The annotation corresponds to the output of the former, extraction and analysis, and the input of the latter, rephrasing and sentence generation. Using terms of information extraction, named entity extraction and temporal expression analysis are completed by this annotation. Therefore, for researchers interested in important sentence extraction or language processing on named entity extraction and temporal expression analysis, the annotation can be referred to as the correct result of their process and used as training data in their machine learning approach. For researchers interested in rephrasing, sentence generation, and information visualization, the annotation can be used as input data in which several fundamental analyses are already completed. In extreme cases, studies on information visualization from text could be conducted without text processing. In this sense, the annotated articles behave as a hub of a multi-model summarization.

# 4 Researches Conducted

Indispensable component technologies of multimodal summarization include extraction of information on statistics from texts as materials of trend information; visualization of statistical information extracted and/or collected; and generation of text that explains statistical information. Many of the researches conducted by MuST participants so far are categorized into these categories, and studies on the synthetic use and fusion of information from multiple sources have been started also. Some researches with a scope beyond the primary MuST framework have emerged, in which multimodal summarization of a topic given by users is compiled. These are characterized as trend mining, which is a version of text mining. In such studies, attempts are made to find trends, which mean topics in vogue in these studies, factors causing the trends, and relationships between those factors by examining patterns of appearance of various keywords such as names of statistics and events.

This section reviews the researches conducted by MuST participants. Through the explanation, references are indicated using a symbol '#' followed by the number of participants shown in Table 1. All titles and abstracts of their works and most of the whole papers can be found on the MuST home page for those interested [3].[2]

## 4.1 Information Extraction from Texts

Information extraction on a given statistic from newspaper articles is a major research theme of the MuST workshop. Information to be extracted is triplets of the name of a statistic, date, and the value of the statistic on that date, which constitute points plotted on the chart depicting the changes of a given statistic. Quadruplets are extracted if a given statistic has a parameter in addition to date, such as companies of market share statistics. The base technology identifies date expressions and numerical expressions using a named entity extraction mechanism, and then organizes those to construct triplets using pattern matching based on the order of appearance of those expressions [#9,#10,#13]; surrounding keywords and surface expressions [#1]; or dependency structures of the phrases containing those [#2,#5,#20]. Difficulties reside in frequent omissions of date expressions and mingling of descriptions of different statistics in one sentence. Moreover, names of statistics can be expressed in various ways and it is very difficult to identify which part describes a given statistic accurately. This problem is often avoided by using knowledge that relates counter suffixes to statistics [#20] or advance filtering using keywords [#10]. In this regard, Mori et al. investigated the components of statistical names and their characteristics, gave those components a precise definition, and then tried to identify them using a machine learning technique [#17].

The distinctive requirement for trend information is that as many pieces of information as possible on the same statistic are expected to be extracted from a series of multiple documents. To achieve this, many researchers are trying to use comparative expressions such as "Oil prices as of this February have dropped 40% since last October" as the source of information extraction and to increase the number of triplets extracted by using some calculations [#1,#19,#20]. Nanba et al. focus on rhetorical relations between sentences crossing multiple documents, and find the parts that describe the transition of the same kinds of information [#10]. Kato et al. consider qualitative expressions on changes such as "increases" and "marks a peak" to be more important for trend information and are trying to extract such information [#19].

A statistic to be processed is not always given, but determination of an important statistic of a given set of documents has been examined. Murata et al. take as an input a set of documents that concern the same topic, identify an item expression by using termfrequency related measures, which corresponds to the name of a statistic, and a counter suffix that plays a major role in those documents, and then extract information related to those. Time-series data of the major statistic can be obtained by extracting date expressions related to the major item expression identified and numerical expressions with the major counter suffix identified. It is also possible to extract a mutual relationship of two major statistics by extracting two kinds of numerical expressions with different major counter suffixes identified, which are related to the major item expression [#4].

## 4.2 Information Visualization

Although most of the systems described above have a mechanism for depicting their extraction results, their visualization method is straightforward since the systems consider only simple time-series data without extra parameters and such data is usually represented using simple line charts. There are a few exceptions in which an appropriate chart type, such as a pie chart and column chart, is selected using knowledge on the characteristics of statistics [#13] or heuristics based on variations of extracted triplets and quadruplets [#9]. Kato et al. visualize extracted qualitative information on changes by placing on a chart a symbolic object with a schematic shape corresponding to the type of changes [#19].

Some studies focus on information visualization. Takama et al. visualize trends of earthquakes using two information graphics: the one relates earthquakes to

[2] all materials are in Japanese.

the places on a topographical map where they occur and the other places them on a time line when they occur [#8]. These information graphics allow users to see earthquakes from different viewpoints and to access their detailed information interactively. Sasakura et al. are addressing the representation of spatial statistical information such as land prices and precipitation, and in order to resolve the problem that such statistics do not always have values for every point in a region of interest, have contrived a two-dimensional representation in which statistical values are represented by color tones that are graded toward the periphery, allowing users to estimate missing values [#6].

Visualization of trend information often involves designing an interface for information access, as is the case with Takama et al.'s study described above. Such an interface allows users to grasp an abstract and overall picture through a visual representation such as information graphics, and through that representation to access more detailed individual information of interest. Matsushita et al. devised a widely applicable interface for accessing information related to time-series statistics [#19]. The interface system organizes documents into a line chart of a given statistic according to the time point that the documents are discussing, and allows users to grasp an abstract through interactive direct operations such as zooming, and to access seamlessly detailed individual information. Itoh et al. are attempting to apply matrix-based visualization based on hierarchical clustering to trend analysis [#12].

In order to generate multimedia presentation in which linguistic and non-linguistic information works cooperatively, Kobayashi et al. are working to generate summaries of multiple documents that explain statistical data coordinately with a given chart of it [#14]. Depending on whether the chart has a fine granularity depicting detailed changes or a coarse granularity showing the outline, the size of document sets is varied for extracting important sentences. Yamamoto et al. are studying how to attach linguistic annotations appropriately to time-series data depicted as a line chart [#11]. An experiment on cabinet approval ratings annotated by factors that are considered to affect such ratings showed that annotations are expected to lie on points with large changes, minimum and maximum points, and so on.

### 4.3 Text Generation for Explaining Time-Series Data

The generation of text that explains statistical information is complementary to the extraction of statistical information from text. These techniques, when combined, allow bidirectional conversion of linguistic and non-linguistic information. Kobayashi et al., using stock price data as an example, approximate the appearance of a given series of time-series data by a fifth-order polynomial curve using a least squares error method, assign linguistic expressions corresponding to the whole and portions of the obtained curve using a dictionary constructed by examining real explanation texts, and then combine those expressions to generate the whole explanation [#14]. Umano et al. divide time-series data into some intervals, assign a linguistic expression to each interval by referring to the global change of data during the interval and its local fluctuation based on a fuzzy set theory, and then compose a total explanation [#7]. Nojima considers that the explanation of time-series data can be represented as a set of rules each of which specifies the value of data under given conditions. By using the complexity of each rule, the number of rules, and the preciseness of explanation generated from the rule set as objective functions, the genetic algorithm totally optimizes those objective functions in order to attain a concise and precise explanation [#3].

### 4.4 Using Multiple Information Sources

Integrating information from multiple sources is important for compiling trend information. Beside the divergence between linguistic and non-linguistic information, there are several genres and several styles even in the same linguistic information. Using two different newspapers, Suzuki et al. examined how many articles were related to each other and how to identify those articles having similar content [#20]. Imaoka et al. examined inter-newspaper applicability of patterns for extracting statistical information, that is, how effectively patterns constructed from newspaper articles work when applied to other newspaper articles [#1]. Meanwhile, Okuda et al. are trying to extract trend information from blog texts [#10]. They found that in blog texts – at least those on subjects concerning economic situations like gasoline prices and product shipment volumes – there are fewer casual and colloquial expressions than expected, and that there are many quotations from newspaper articles. Encouraged by these findings, they examined and showed that an information extraction system designed for newspaper articles can be easily transported to blog texts and deliver high performance.

### 4.5 Trend Mining

Under the research framework of the MuST workshop, a system that compiles trend information of a given topic is expected to be constructed, but some researches have started going beyond this framework. Examples include a kind of text mining that tries to find hidden relationships among several topics and to understand their meanings and implications. Terachi et al. are applying a method for market analysis in order to analyze patterns of keyword appearance us-

ing OLAP, a multi-dimensional data analysis system [#15]. Employing this framework, they expect to find what keyword or topic is in vogue and which will become popular next. Kitamura et al. are challenging the topic detection and tracking (TDT) task on the MuST data set, and examined whether it is possible to trace a series of articles concerning the trend of a specific topic [#18]. Yamamoto et al. showed that keywords in related articles appearing frequently in a given time period are useful for suggesting factors that affect changes in some statistic such as cabinet approval ratings during the same period. They developed such a method to find hidden relationships among topics represented as the names of statistics and event names by examining temporal co-occurrence patterns thereof [#11]. Saito et al. have proposed a method to automatically identify and collect statistical terms, which are closely related to names of statistics, and are trying to explore causal relationship networks of those automatically identified terms by calculating their spatial co-occurrence, that is, collocation pattern [#21].

## 5 Evaluation of the Workshop Framework

As shown in the review of the previous session, several research activities are ongoing in the MuST workshop. The scope of such work has now gone beyond the initial framework in which linguistic and non-linguistic information is cooperatively processed and used for compiling trend information, and has extended to trend mining in which hidden trends can be found and analyzed. Although such studies are diverging from conventional research fields, researchers have gathered with a common interest in trend information and created a field of intensive discussion, which is a remarkable achievement of the MuST workshop. The researchers share a common recognition on what is trend information and what is related to it by reviewing materials in the MuST data set, and this shared recognition has accelerated their discussions. The MuST data set is thus playing a crucial role for establishing this shared recognition.

On the other hand, it is still doubtful whether the annotation of the MuST data set has been used as expected in their researches. Although the annotation of the data set was expected to be the intermediate results of text summarization and information extraction, most researches on information extraction are using the output of a general-purpose named entity extraction system while a few presume the annotation in the data set and develop mechanisms based on it. Use of selected topics and articles in the data set is prevailing, however, showing that the MuST data set is a useful shared material to be examined and processed, though some researches need more articles than are available in the data set. It is thus concluded that the role as the shared material is important, after the main contribution of establishing a shared recognition on trend information. In addition, some researches on information visualization obtained data to be visualized from the data set using a XML parser, and the data set appears to be used frequently for preliminary analysis at the initial stage of research and for making references and correct data for evaluation. The annotation itself has thus been of some use.

While the first purpose of the workshop, which is to promote discussion and community conformity, has been achieved successfully, the second one, which is to construct and accumulate resources such as tools and corpora, remains to be achieved. A large amount of data has been collected and shared, but has not yet been used much. More active use is expected as the quality and amount are improved through further research, and the data will then be useful for researchers starting research on related topics in future.

## 6 Future Direction

The MuST workshop has been a decided success. Encouraged by this success, we will continue to hold the current course and start the third cycle, which includes the final workshop scheduled for March 2008. We also have proposed that the MuST workshop be an official task of the NTCIR-7 workshop. If this proposal is accepted, the third cycle will be extended according to the schedule of the entire NTCIR-7 workshop, and all research results during the third cycle and consecutive half year will be presented at the NTCIR-7 workshop meeting scheduled to be held in December 2008. But even if it is accepted as an official task of the NTCIR, the fundamental part of the MuST workshop will not change; its main objective remains to encourage cooperative and competitive studies using shared research resources rather than objective evaluation conducted in an evaluation workshop. It should be considered, however, that as some shared research themes such as information extraction on statistical information have emerged, objective evaluation of those themes could be introduced as a part of the workshop.

The present data set is large enough for its current usage and so we do not plan to extend it, and we have also frozen the annotation criteria to the present ones. Remaining tasks include fixing some errors and incoherencies in the current version, and some specifications might be modified or elaborated during such correction process. If some objective evaluation is conducted in the workshop, a test set for that purpose will be needed, which would be constructed according to the current criteria of the annotation.

In order to accelerate researches on the synthetic use of information from different sources, which is attracting much research interest, it is important to collect non-linguistic information such as charts and nu-

merical data corresponding to the topics currently covered by the articles in the data set and texts in different genres such as blogs parallel to those. However, there is a serious problem that the collection of current articles was published too long ago to collect such materials, and this problem needs to be addressed as soon as possible.

The target of multimodal summarization is not limited to trend information. Emphasizing this point and also emphasizing the importance of an information access environment that seamlessly assists reviewing and accessing information, a novel framework named *information compilation* has been proposed [2]. The developments in this new framework will be examined through the third cycle of the MuST workshop and beyond.

## 7 Conclusion

This paper reported on the MuST workshop, which was designed to encourage cooperative and competitive studies on multimodal summarization for trend information. It has been run based on approximately one year as one cycle, and it finished its second cycle in March 2007. Several researches with a broader scope than initially expected have been conducted in the workshop. The workshop has successfully accomplished its mission of promoting discussion and community conformity by sharing the same research resource and by conducting research concerning a common theme in a broad sense. In the coming third cycle, continuing to hold the current course, we will also focus on constructing and accumulating resources such as tools and corpora and on establishing shared objective evaluation measures.

## Acknowledgments

The participants of the MuST workshop inspired many ideas and insights, for which we are grateful. The MuST workshop and related research are supported in part by grants from a joint research project of Nippon Telegraph and Telephone Corp. and the University of Tokyo, and a joint research project of the National Institute of Informatics and the University of Tokyo.

## References

[1] Tsuneaki Kato, Mitsunori Matsushita, and Noriko Kand. MuST: A Wroskshop on Multimodal Summarization for Trend Information. *Proceedings of the Fifth NTCIR Workshop Meeting on Evaluation of Information Access Technologies*, pp. 556–563, 2005.

[2] Tsuneaki Kato and Mitsunori Matsushita. Toward Information Compilation (in Japanese). *The 20th Annual Conference of Japanese Society for Artificial Intelligence*, 1D3-2, 2006.

[3] http://must.c.u-tokyo.ac.jp.