Automatic Speech Recognition and Video Retrieval System for the Japanese Diet

Mikitaka Masuyama
Tatsuya Kawahara

July 2019

Automatic Speech Recognition and Video Retrieval System for the Japanese Diet

MASUYAMA Mikitaka

National Graduate Institute for Policy Studies

KAWAHARA Tatsuya

Kyoto University

Summary

The Japanese Diet uses the Automatic Speech Recognition (ASR) system, which directly transcribes parliamentary speech made in both plenary and committee meetings. ASR performance is monitored for most of the meetings, consistently achieving over 90 percent of Japanese character accuracy. By using the ASR for Diet deliberations, an internet video retrieval system is developed to create timestamp data to match the minutes of parliamentary meetings and video feeds. Our video retrieval system allows one to pinpoint and play the parliamentary video clips corresponding to the minutes of meetings by means of keyword-search. In this paper, we offer an overview of the video retrieval system, and suggest a variety of ways in which our video retrieval system can be utilized. By analyzing the differences between faithful transcripts and official minutes, we depart from tradition of focusing on written official minutes and shed new light on the complex and multifaceted nature of parliamentary discussion.

Introduction

Seeking innovative ways to provide universal access to policy information, we have launched an internet video retrieval system for the Japanese Diet.[1] Although many parliaments offer online video streaming of speech and debate, those video library systems merely provide a search engine for retrieval of videos by date and by parliamentary meeting theme. However, even if we successfully retrieve the parliamentary video we want, we need to watch the video from the beginning to the point in the speech or debate we are particularly interested in. Using the latest sound recognition techniques to create timestamp data to match parliamentary video feeds and the minutes of proceedings, we have developed an internet video retrieval system for Diet deliberations, in which one can pinpoint and play the parliamentary video clips corresponding to the minutes of proceedings by means of keyword search.

With our video retrieval system one can directly retrieve the video feed segment one is particularly interested in, gain a visual understanding of the flow of parliamentary debate, and check the facial expressions and body language of the speaker. Our system captions the videos, so it offers an alternative means of accessing parliamentary deliberation clips for those who have visual and hearing impairments. Since it is easy to share the URL identifying a moment in a video feed via SNS, our system has great potential to boost the usage of Diet deliberation videos by researchers and ordinary citizens as well.

In this paper, we offer an overview of the video retrieval system we have developed, and demonstrate how one can retrieve video streaming on user terminals that do not support Japanese language input. We also suggest a variety of ways in which our video retrieval system can be utilized, and report a first systematic analysis of the mismatches between the official minutes and the results of speech recognition of recordings of parliamentary meetings.

Video Retrieval System for Diet Deliberations

The Japanese Constitution stipulates that each house of the Diet shall keep a record of proceedings and put it into general circulation. Since the opening of the Imperial Diet in 1890, verbatim records had been made by manual shorthand. However, early in this century, the government terminated recruiting stenographers, and investigated alternative methods for transcribing parliamentary speeches.

Kawahara, one of the co-authors of this paper, has developed the automatic speech recognition (ASR) technology, which has been deployed in the transcription system for the House of Representatives, one of the two houses of the Diet. To achieve high recognition performance in spontaneous meeting speech, Kawahara has investigated an efficient training scheme with minimal supervision that can exploit a huge amount of real data, and proposed a lightly-supervised training scheme based on statistical language model transformation, which fills the gap between faithful transcripts of spoken utterances and final texts for documentation. Once the mapping is trained, faithful transcripts for training both acoustic and language models are no longer needed. The ASR system has been consistently achieved character accuracy of over 90% since 2010, which is useful for streamlining the transcription process.

In the survey conducted in the House, a majority of the reporters felt that it took less time and labor to finish a draft with the ASR system, and more than 80% said they

---

[1] Masuyama (2016a, 2016b), Masuyama and Takeda (2014, 2015).

are satisfied with the performance of the ASR system. Some also expressed the positive opinion that the system would make it possible for those who have not been trained on stenography to produce an edited draft upon proper training. Moreover, we observed a significant decrease in the edits from the actual utterance, for example changing a colloquial expression to formal expression, in making the transcript over the past 10 years. This probably has to do with the increasing availability of SNS and real-time video streaming of the meeting. Another possible factor is the use of the ASR system which generates faithful transcript.[2]

The Diet Library currently provides via internet the digitized minutes of parliamentary meetings. Although those are not considered as "official" records, they are amenable to keyword searching. On the other hand, we can watch the online live streaming of proceedings at the secretariat website of each house. We can also search the video library and watch video streaming of parliamentary proceedings. Both houses originally made video streaming of plenary and committee meetings available for only one year, but the House of Representatives alone changed its policy so that the videos of proceedings since 2010 are currently available for viewing.

http://www.shugiintv.go.jp/index.php

http://www.webtv.sangiin.go.jp/webtv/index.php

Diet deliberation videos can be searched by meeting date, meeting title, subject and speaker, although only the first two search options are offered in the English interface. However, even if we successfully retrieve the desired Diet deliberation video, we need to watch the video streaming from the beginning to the speech or debate segment we are particularly interested in. It is not uncommon for a committee meeting to last more than 7 hours. While the breakdown of the video by questioner is available in the Japanese interface, streamed video segments are usually 30 to 60 minutes long. No such breakdown is available in the English interface. Moreover, replies to parliamentary questions are included in the video, arranged by questioner. Thus, we have no way to search the Diet deliberation videos of prime ministers and cabinet ministers answering parliamentary questions.

By linking the Diet Library's proceedings database and the Diet secretariats' deliberation video libraries, our "Video Retrieval System for Diet Deliberations" makes it possible to retrieve the deliberation video clips corresponding to the minutes of the proceedings through keyword searching:

http://gclip1.grips.ac.jp/video/

For instance, our system allows us to pinpoint and play the related video clip when we see an item such as "MP X made remark Y during parliamentary debate." With our system we can directly retrieve the portion of the video feed we are particularly interested in, gain a visual understanding of the flow of parliamentary debate and check the facial expressions and body language of the speaker, all of which are not possible from a simple reading of the minutes of parliamentary proceedings.
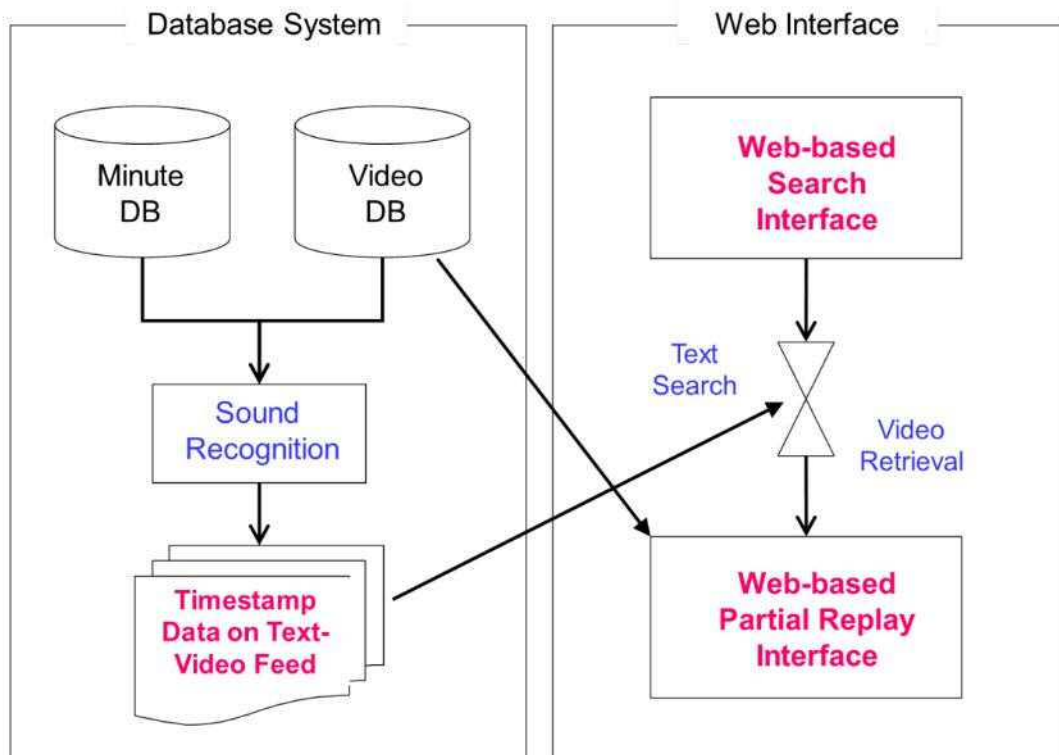
Unlike the Diet secretariats' websites, our system creates and adds subtitles to the Diet deliberation videos, and thus offers those suffering from hearing impairment a means of accessing the deliberation video library. At present the Diet Library's website is not

---

[2] See Kawahara (2012 and 2017) and Akita and Kawahara (2013) for the details. The House of Councillors Secretariat relies solely on stenographic transcripts.

compatible with the screen readers used to assist blind and visually impaired users. Even if we can successfully search the proceedings database and reach the speech or debate of particular interest, we need to use text vocalizing software that produces synthetic voices which do not resemble that of the original speaker. In contrast, our video retrieval system allows us to see instantly video of what was actually spoken in the Diet.

Technically speaking, our video retrieval system consists of two sub-systems. As illustrated in Figure 1, one of the sub-systems uses the latest sound recognition techniques to create timestamp data to match the Diet Library's proceedings database (Minute DB) and the Diet secretariats' deliberation video databases (Video DB). The second sub-system uses the timestamp data to search the Diet proceedings and retrieve the Diet deliberation videos corresponding to the minutes of the proceedings by means of keyword search (Web-based Search Interface). The results of keyword searches are deliberation video links, and the portion of video we are particularly interested in can be played partially by clicking the URL link for the deliberation video stored in the Diet secretariats' databases.

Figure 1: Process for Linking Diet Proceedings and Deliberation Videos

"Video Retrieval System for Diet Deliberations" has been in operation and publicly available since November 2012. It is possible to keyword search all the plenary and committee meetings in the House of Representatives since January 2010 and those in the House of Councillors since December 2012. Below, we briefly describe how our video retrieval system works. Figure 2 shows the top page of our web-based search interface that allows us to search for deliberation video segments by typing keywords. The Japanese interface will appear when the user clicks on "Japanese" in the upper right hand corner.
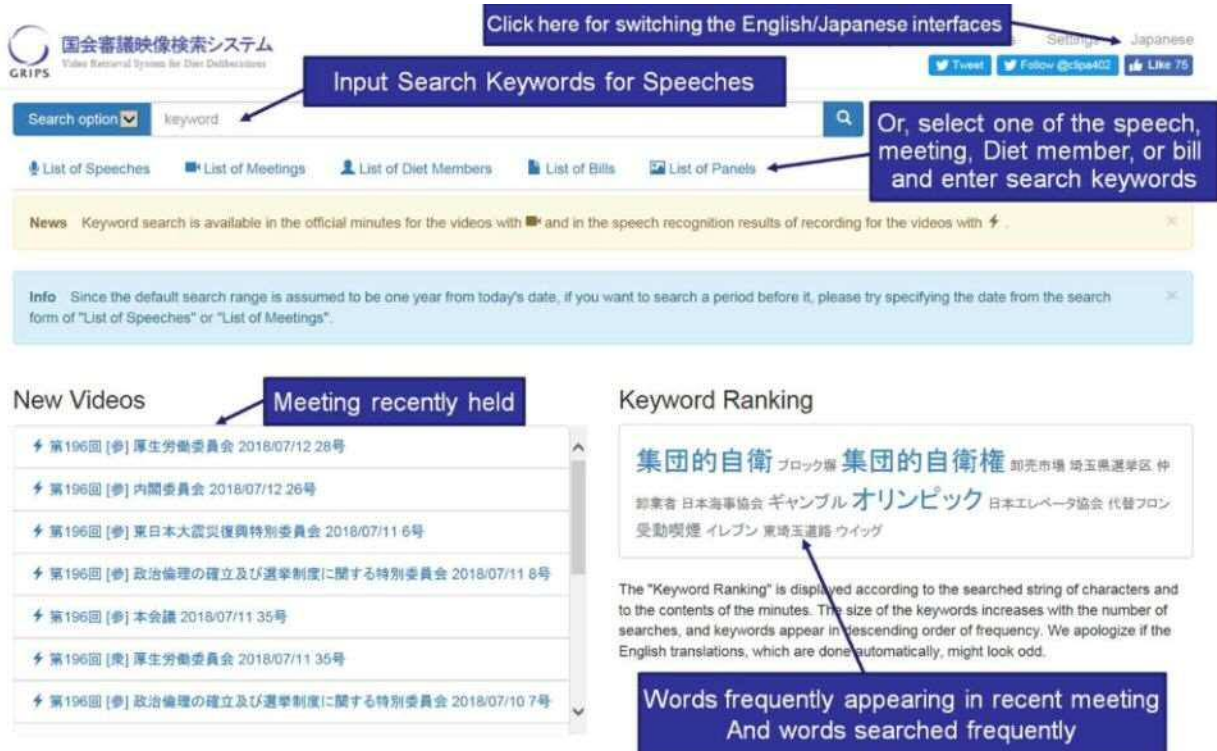
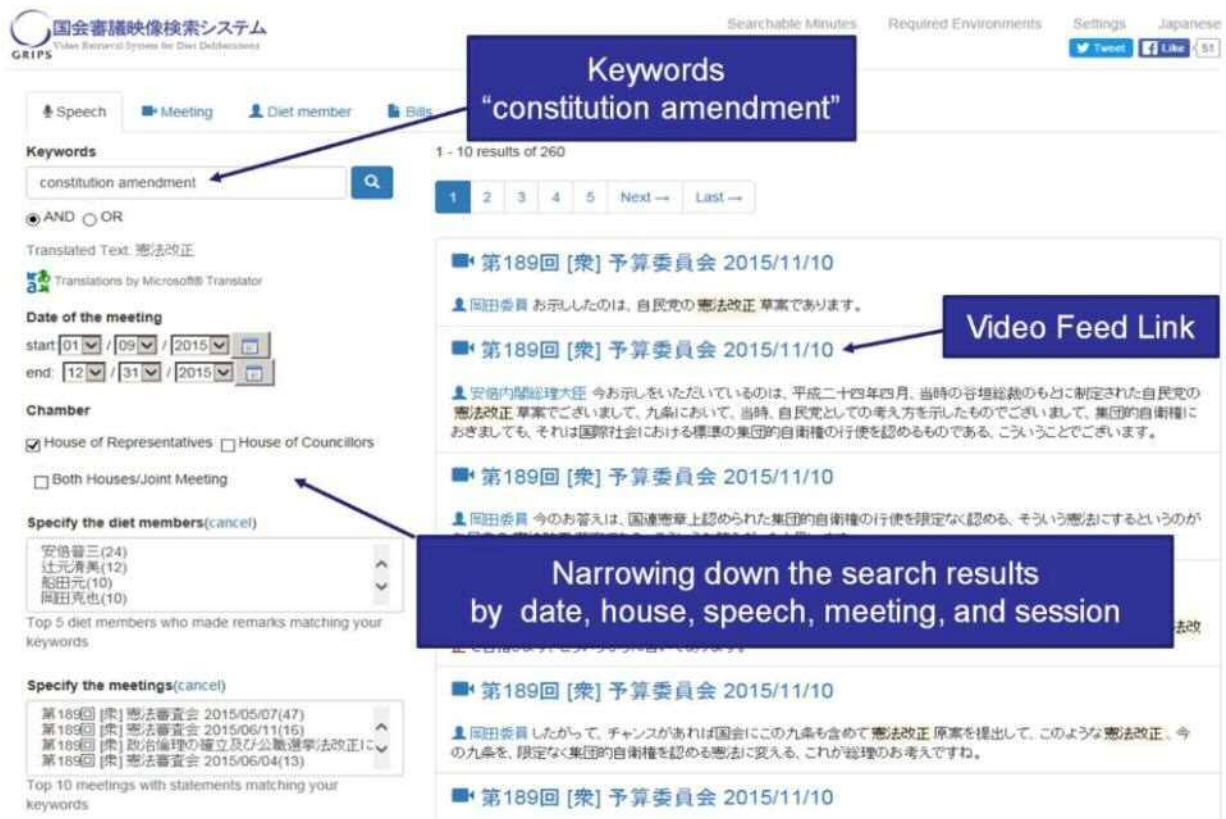Figure 2: Keyword Search Interface in English

Figure 3: Video Feed Link

As shown in Figure 2, one can type English keywords separated by spaces in the

search field, and they will be translated automatically into Japanese and used in keyword searching. Any combination of keywords can be used. For instance, one could type "Abe Shinzo" (the name of the current Prime Minister of Japan) and "constitution amendment" in the search field. Alternatively, one could first select one of the four categories (speech, meeting, Diet member, and bill) from "search option" and type keywords in the search field. If one selects "Diet member" from "search option" and types "Abe Shinzo constitution amendment" in the search field, each of the keywords will be first considered as the name of a person and then as either a word in a speech, the title of a meeting, or the agenda of a meeting, giving priority to search results where Abe Shinzo uttered 'constitution' and 'amendment' in his speech, over results where some Diet members uttered 'Abe' 'Shinzo' 'constitution' and 'amendment' in their speeches.

The search results are listed as URL links. Once we hit the search button a screen similar to Figure 3 is displayed, listing the search results in order of date (keywords "constitution amendment" in the House of Representatives for the period September through December 2015). In addition to specifying AND/OR search options by selecting one of the boxes under the search field, we can filter the search results by date and limit the search to one or both of the two houses, or to joint meetings of the two houses. As shown in Figure 3, additional search options include speakers (top 5) and meetings (top 10) with words with utterances matching the search text, and allow us to further narrow the search results by selecting one of the meetings and one of the speakers.

Figure 4: Video Replay Interface

If we click one of the video feed links, the screen content will be similar to that in Figure 4. Subtitles are shown under the video. The speech list is shown on the right side, and speech currently playing is highlighted. The video will play for either one minute or for three speeches. Alternatively, we can keep playing the video by clicking the play button in the toolbar under the video. By double-clicking any speech in the list on the right side, we can instantly watch the video stream of the speeches before and after the speech found by keyword search. Once the user has moved on to another speech, the original speech found by keyword search remains highlighted in yellow.

Further, the URL for the corresponding segment of video streaming is shown below the video, and we can easily share the URL via SNS by clicking the tweet button while the video stream is playing. The text of the speech and the URL will immediately appear in the tweet box after the tweet button is clicked. At the bottom of the page the profile of the speaker is provided, followed by a list of agendas and a list of the Diet members attending the meeting (not shown in Figure 4)

To assist keyword searching, "Keyword Ranking" on the right side of the top page lists 15 words uttered in the Diet proceedings, in descending order of frequency, placing more weight on frequency in later parliamentary meetings than earlier and more weight on single meetings than multiple meetings. The font-size of the words increases with the number of searches, reflecting the attention given to Diet deliberation videos. An English translation pops up when the mouse hovers over any word.

Usage beyond Keyword Search

Our video retrieval system can be utilized in a variety of ways. For instance, we can create a list of search results with the query "constitution amendment" and the name of the speaker, "Abe Shinzo." By clicking one of the video links in the list we can instantly retrieve video of the speeches by Prime Minister Abe Shinzo, with his own voice, facial expressions and body language, where he mentioned constitutional amendment.

Diet members are increasingly posting information about their activities on the websites. Some use their websites to display the minutes of parliamentary proceedings, and some even edit and upload deliberation videos on their websites. In contrast, our video retrieval system allows us to obtain the URL for a moment of video streaming and to create a list of video links without downloading and editing the video files.

Furthermore, we can use Twitter to create a list of parliamentary speeches.[3] For instance, prime ministers customarily begin answering questions in plenary meetings by saying "There is a question regarding X." Thus we can narrow the above-mentioned search results to speeches beginning with "There is a question regarding constitutional amendment," and tweet the speeches and their video links to create a list of Prime Minister Abe's plenary speeches on the issue of constitutional amendment.

Another way of utilizing the interfaces for keyword searching and partial replay is to post deliberation video links to internet news.

> *On February 13, 2014, Abe Shinzo told the Budget Committee of the House of Representatives that he is the "ultimate arbiter" of affairs concerning the constitutional interpretation.*

---

[3] Masuyama (2017b).

For instance, if we find a report like the above in an internet newspaper, featuring the remarks made by Prime Minister Abe in the Diet, we can enhance the internet news visually by using our video retrieval system and inserting the video link (below) for the moment of video streaming in question. Clicking on the link will result in the instant playing of the video of the moment of Prime Minister Abe's controversial remark (Figure 5).

http://gclip1.grips.ac.jp/video/video/1117?t=45m18s

国会審議映像検索システム　Video Retrieval System for Diet Deliberations

第186回 [衆] 予算委員会 2014/02/12

[安倍内閣総理大臣] 先ほど来、法制局長官の答弁を求めていますが、最高の責任者は私です。

Figure 5: Abe Shinzo on Constitutional Interpretation

The minutes of the proceedings are an important source of the content of discussion in the Diet, but they are not the whole story. For instance, supplementary materials are often used in parliamentary meetings, and discussions refer extensively to graphic materials such as figures and tables. Such supplementary materials are not included as part of the minutes unless a Diet member attending the meeting requests that they be included. Although the secretariats and the library of the Diet keep the supplementary materials used in parliamentary meetings, the existence of those materials is not widely known and they are hardly used.

We are working on combining sound and pattern recognition techniques to distinguish between the portions of videos that do or do not focus on the speaker and automatically extract video clips including the moments focusing on supplementary materials used in committee meetings. Moreover, the minutes are silent with regard to non-verbal communication, and we are developing a web-based program to automatically extract and analyze facial expressions and body language of the speaker. Figure 6 illustrates the supplementary materials found by means of keyword searching and pattern recognition, compiled into the database using text recognition in video images. Figure 7 shows an example of the results of emotion recognition and analysis.
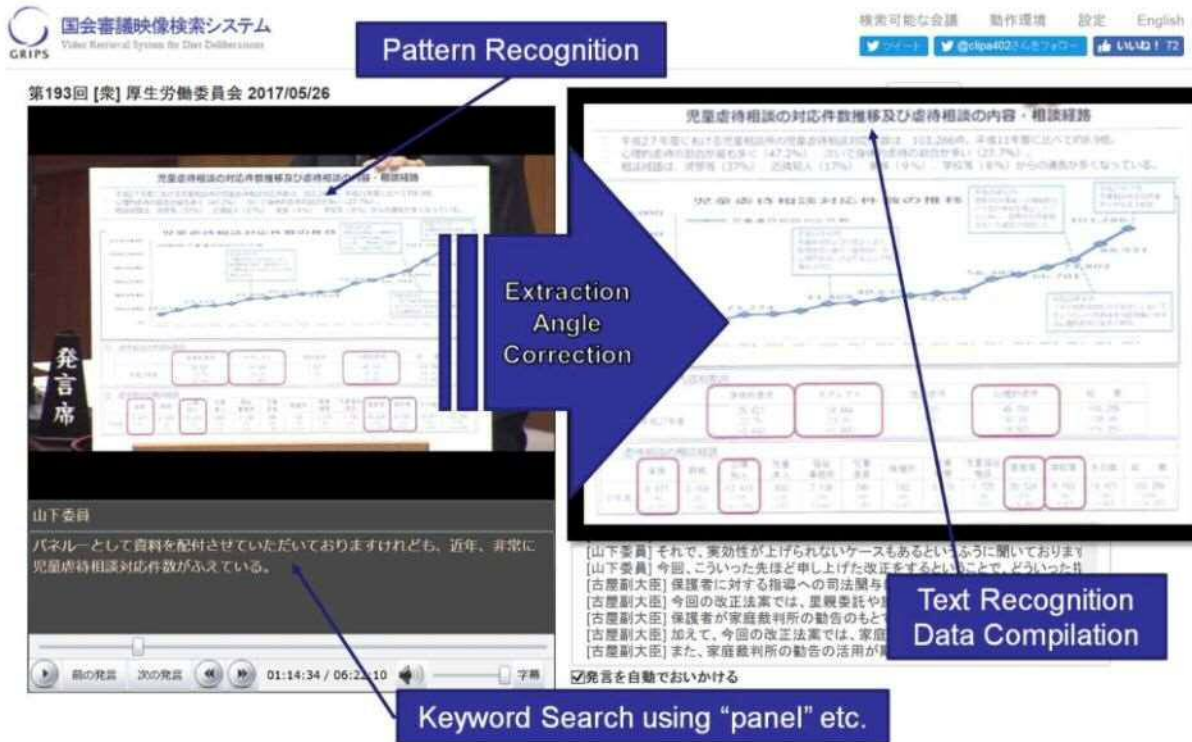
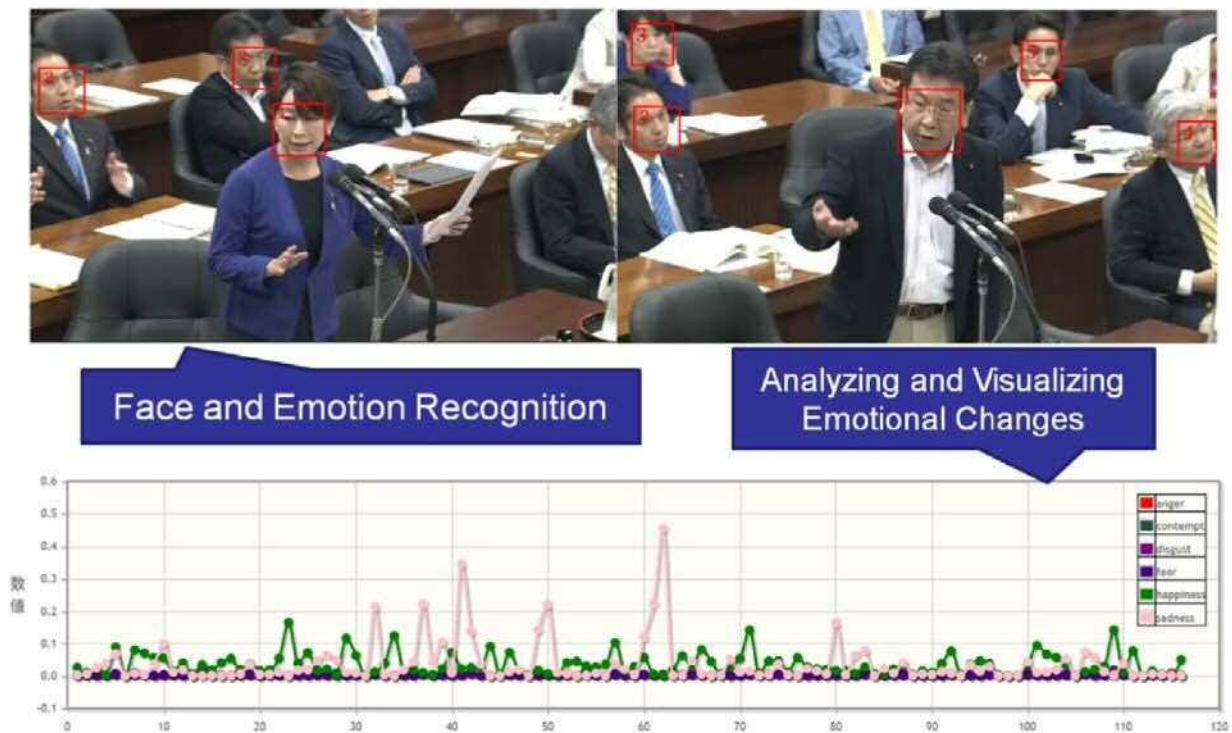Figure 6: Supplementary Materials Found by Keyword Searching

Figure 7: An Example of Emotion Recognition and Analysis

By taking advantage of our system's combining of the interfaces for keyword searching and partial replay, we may be able to conduct experimental studies to

determine how visual information affects the understanding of Diet deliberations.[4] We can think of the informatics of Diet deliberations as a multidisciplinary scientific approach to parliamentary deliberations, involving analysis not only of textual information, but also of audio-visual information.[5]

A unique aspect of our video retrieval system is that we use sound recognition techniques to create timestamp data to match Diet proceedings and deliberation videos. In other words, we deal with two types of text information related to parliamentary meetings. Both are written versions of speeches made in the Diet. The information derived from speech recognition is "correct" in the sense that it captures 100 percent of what was actually spoken in the Diet, although it may contain irrelevant filler and words wrongly recognized due to both individual speaker factors such as intonation and pronunciation and technical and environmental factors such as recording quality and noise. On the other hand, the proceedings of the Diet meetings become "official" after transcription to eliminate filler, correct inappropriate wording, and add commas and periods so that the speech in the Diet can make sense as a written language.

With a web-based program to automatically calculate correspondence rates as part of the standard procedure for creating timestamp data to match proceedings and deliberation videos (Figure 8), we can systematically analyze the correspondence between the official minutes and speech recognition results by meeting, speaker, etc. in cross-sectional and longitudinal manners.

As a first systematic analysis of the mismatches between the official minutes and the results of speech recognition of parliamentary videos, we compiled data on the words in the minutes, which differ from the speech recognition results for the entire year of 2018, covering both the 196th ordinary session and the 197th extraordinary session.[6] To see what type of words are not accurately recognized, we analyze the frequency of such words in the data set containing 97,663 speeches in total (more than 40 million Japanese characters in total), and visualize the results by using word cloud techniques (Figure 9). As the figure illustrates, the words the ASR has difficulty in recognizing sounds accurately are the proper nouns, typically including names of persons ("*Kake* Academy" and "*Moritomo* Academy"), and the common nouns related to the topics of parliamentary meetings, which may be somewhat peculiar from the ordinary linguistic point of view. Examples are "regional revitalization" "passive smoking" "death from overwork" "non-fulltime employment" (respectively "*Chihô Sôsei*" "*Judô kitsuen*" "*Karôshi*" "*Hiseiki*" in Japanese).

---

[4] Masuyama (2017a, 2018a, 2018b and 2018c).

[5] Our video retrieval system offers a particularly useful tool for scholars interested in the differences between spoken and written language. Matsuda (2016) used our system to check whether Diet members say "Ba'ai" or "Bayai," which are written with the same Japanese characters. The minutes of parliamentary proceedings do not enable us to check if the two words are pronounced differently.

[6] There are three categories of Diet sessions: ordinary, extraordinary, and special sessions. An ordinary session is convened once per year to deliberate the annual budget and related legislation, for a term of 150 days by law, subject to extension or early termination. Currently, an ordinary session is convened in January. The cabinet may decide to convoke an extraordinary session of the Diet whenever it is deemed necessary or at the request of a quarter of the total members of either house. A special session must be convoked within thirty days of a general election that is called within forty days from the date of the House of Representatives dissolution.

Speaker | Speech Recognition | Official Minute | Difference | Correspondence Rate

会議名	発言者	発言内容（音声認識）	発言内容	差分分析（削除、挿入、一致）	正文率
第189回 [衆] 予算委員会 2015-03-02	山尾委員	山岡君民主党の女性で地盤は山をしておりです。きょうは私野党としての初質問でもございます。その中で大きく分けて二つまあやはり政治と金の問題についてまだ聞きたいれていない点そしてまたで公園では主に有村大臣と子供女性家族について建設的な議論したいと思っております。よろしくお願いいたします。最初に望月大臣に伺いました。この週末新たに報道がございました。二千十一年からミス十二年にかけての続い様からの献金も政治資金規正法二十二の三に違反する疑いがある事実関係確認をさせてください。環境省から一般社団法人を通じて二千十一年九月に三十五万円の補助金同年十二月に三十三万円の補助金が珠洲四区に交付決定をされこの一年内十一年十一月に百四十万円の献金十二年一月に三百九十万円の献金がこの珠洲様から望月大臣の静岡四区総支部に向かってなされこの事実関係に間違いはございます。望月大臣	民主党、女性で二番バッターの山尾志桜里です。きょうは、私、野党としての初質問でもございます。そんな中で、大きく分けて二つ、やはり政治と金の問題について、まだ聞き足りていない点、そしてまた後段では、主に有村大臣と子供、女性、家族について建設的な議論をしたいと思っております。よろしくお願いいたします。最初に、望月大臣に伺います。この週末、新たに報道がございました。二〇一一年から二〇一二年にかけての鈴与からの献金も政治資金規正法二十二の三に違反する疑いがあると、事実関係の確認をさせてください。環境省から一般社団法人を通じて二〇一一年九月に三十五万円の補助金、同年十二月に三十三万円の補助金が鈴与に交付決定をされた。そして、この一年内、一一年十一月に百四十万円の献金、一二年十月に三百九十万円の献金がこの鈴与から望月大臣の静岡四区総支部に向かってなされた。この事実関係に間違いはございませんか。	山岡君民主党の女性で地盤は山をしており、女性で二番バッターの山尾志桜里です。きょうは、私、野党としての初質問でもございます。そのんな中で、大きく分けて二つまあ、やはり政治と金の問題について、まだ聞きたいれ足りていない点、そしてまたで公園後段では、主に有村大臣と子供、女性、家族について建設的な議論をしたいと思っております。よろしくお願いいたします。最初に、望月大臣に伺います。この週末、新たに報道がございました。二千十〇一一年からミス十二〇一二年にかけての続い様鈴与からの献金も政治資金規正法二十二の三に違反する疑いがあると、事実関係の確認をさせてください。環境省から一般社団法人を通じて二千十〇一一年九月に三十五万円の補助金、同年十二月に三十三万円の補助金が珠洲四鈴与に交付決定をされた。そして、この一年内、一一年十一月に百四十万円の献金十二年一、一二年十月に三百九十万円の献金がこの珠洲様鈴与から望月大臣の静岡四区総支部に向かってなされた。この事実関係に間違いはございます。望月大臣せんか。	83.59%

XXAXX | XXBXX | XXABXX

Figure 8: How Minutes and Speech Recognition Results Differ

Figure 9: Frequently Misrecognized Words

One of the common nouns most frequently misrecognized is "committee chair"

("*Iinchô*"). Such common nouns to describe the titles of persons and positions are usually mentioned with names, and frequently considered as misrecognized in combination with the names of persons although the common nouns themselves are accurately recognized. For instance, "Chairman Kim Jong-un," the title and name of the North Korea's leader, is mentioned frequently during the parliamentary discussions, and the "chairman" part is "*Iinchô*" in Japanese. A foreign person's name, which often needs to be rendered in a phonetical alphabet (*Katakana*), is not accurately recognized and thus the accompanying noun is considered as misrecognized even though the title itself is properly recognized.

Moreover, the titles of parliamentary actors such as "committee chair" ("*Iinchô*") and "witness" ("*Sankônin*") as well as parliamentary organizations such as "committee" ("*Iinkai*") and "committee directorate" ("*Rijikai*") are often mentioned at the beginning and closing of speeches, usually requesting either permission to speak or an action from the person or organization being addressed. The reason why these words are frequently misrecognized may have to do with the fact that it is relatively challenging for speech recognition to properly detect the beginning and end points of an utterance. Like "committee chair" ("*Iinchô*"), it may also be challenging to accurately recognize the character strings that either include a double consonant or a contracted sound. Another example is "passive smoking" ("*Judô kitsuen*"). At the same time, there are some frequently misrecognized common nouns without a double consonant or a contracted sound.

Table 1 categorizes the type of words misrecognized more than 50 times during 2018 (the frequency ranging from the maximum of 882 times to the minimum of 53 times). Table 2 lists the frequency of the common nouns in the other category without a double consonant or a contracted sound. For the frequently misrecognized common nouns without a double consonant and a contracted sound, we have checked the 21 actual utterances mentioning "Self Defense Force" ("*Jieitai*") in the House of Representatives meetings during the 197th session. Figure 10 shows examples of the contexts in which the utterances are misrecognized. The URL identifies the moment in deliberation video so that we can easily check how the word is actually uttered (see Appendix for the list of the 21 utterances).

Table 1: Type of Words Misrecognized

Type of Words	With double consonant contracted sound	Without double consonant contracted sound
Proper Noun (Foreign Origin)	13	15
Proper Noun (Japanese)	7	7
Common Noun (Parliamentary)	12	13
Other	16	11
All	48	46

Table 2: Common Words Frequently Misrecognized

English Translation	Japanese Original	#
Regional revitalization	*Chihô Sôsei*	396
Death from overwork	*Karôshi*	318
Non-fulltime employment	*Hiseiki*	284
Virtual currency	*Kasô tsûka*	131
Self Defense Force	*Jieitai*	127
Law revision	*Hôkaisei*	113
Common-law marriage	*Jijitsukon*	108
Eastern earthquake	*Tôbu Jishin*	74
Sense of crisis	*Kikikan*	69
Concrete	*Gutaiteki*	63
Foreigner	*Gaikokukjin*	62

Meeting	Budget Committee 2018/11/1
Speaker	Minister Iwaya
Match Minute Sound	そのような話は、実は私も防衛大臣に就任する前に、地元の自衛隊全体家族会の皆さんから聞かされたこと事がございました。 http://gclip1.grips.ac.jp/video/video/7589?t=8h8m3s
Match Minute Sound	十一月十四八日に開催した平成三十年度自衛隊絶対記念日観閲式に御出席いただいた国会議員につきましては、政党別に申し上げますと、。自由民主党五名、米公明党四名、立憲民主党三名、計経営十二名でございましたす。 http://gclip1.grips.ac.jp/video/video/7589?t=8h9m47s

Figure 10: Examples of the Speech with a Misrecognized Word

As the figure illustrates, "Self Defense Force" ("*Jieitai*") is sometimes misrecognized as "Entire" ("*Zentai*") and "Absolute" ("*Zettai*") when the noun is not followed by a postpositional particle. In these cases, the word is directly followed by another noun "family association" ("*Kazokukai*") and "anniversary" ("*Kinenbi*"). While a two-character string is the most common in Japanese, "Self Defense Force" ("*Jieitai*") is a three-character string and may become further difficult to be accurately recognized when uttered in combination with other nouns. Of the 21 utterances, 11 are such cases that the word is followed by another noun, and 2 are the cases that the word is considered as misrecognized in combination with the following utterance that is not accurately recognized. There is one case that the whole sentence including "Self Defense Force" ("*Jieitai*") is not recognized (2018/10/29 Plenary Prime Minister Abe).

Our analysis of deliberation videos suggests that 5 of the remaining 7 are the cases that the speakers tend to use filler sounds:

http://gclip1.grips.ac.jp/video/video/7605?t=49m29s

http://gclip1.grips.ac.jp/video/video/7605?t=2h38m56s

http://gclip1.grips.ac.jp/video/video/7637?t=45m17s

13

http://gclip1.grips.ac.jp/video/video/7637?t=2h24m14s

http://gclip1.grips.ac.jp/video/video/7714?t=3h32m26s

The following 2 are the cases that the speakers tend to talk in a somewhat casual way:

http://gclip1.grips.ac.jp/video/video/7589?t=8h2m23s

http://gclip1.grips.ac.jp/video/video/7714?t=2h1m47s

Conclusion

Although it is still at the developing stage, our video retrieval system is the world's first attempt to retrieve parliamentary video streaming with a keyword search in the parliamentary minutes, offering alternative means of access to parliamentary videos for those who have visual and hearing impairments. Also, by incorporating automatic translation into the web-based keyword search interface, we have made our system available from computers without Japanese text capability.

As part of the standard procedure for creating timestamp data to match proceedings and deliberation videos, our video retrieval system automatically analyzes the correspondence between the official minutes and speech recognition results. In this paper, we compiled data on the words in the minutes, which differ from the speech recognition results, and analyzed the frequency of such words in the data set of more than 40 million Japanese characters. In addition to proper nouns, we found that the words the ASR has difficulty in recognizing sounds accurately are the common nouns somewhat peculiar from the ordinary linguistic point of view and the character strings that either include a double consonant or a contracted sound. Also, it may be challenging to recognize common nouns when they are not followed by a postpositional particle. Moreover, our analysis of deliberation videos suggests that the speaker's talking style matters, which can be easily checked by using our video retrieval system.

Our video retrieval system has great potential to boost the usage of parliamentary videos in relation to policy-making. The sound recognition techniques for creating timestamp data for matching video and text information are applicable to a wide range of proceedings, including those of local assemblies and administrative councils,[7] as well as other types of videos such as TV news clips. Our system is expected to enhance discussion thanks to a better understanding of policy making in the Japanese Diet, and to indicate a direction for the development of parliamentary video streaming systems in other countries as well. We can think of the informatics of Diet deliberations as a multidisciplinary scientific approach to parliamentary deliberations, involving analysis not only of textual information, but also of audio-visual information. By so doing, we can shed new light on how information technology advancements affect the procedural innovation in which political behavior is translated into structures and rules, and institutions are established and transformed.

---

[7] We have developed a similar video retrieval system for the Japanese local assemblies, and a sample of deliberation videos are searchable by keywords for selected prefectural and municipal assemblies.

http://gclip1.grips.ac.jp/local-assembly/

References

Akita, Yuya and Tatsuya Kawahara. 2013. "Onsei Ninshiki wo Mochiita Onrain Jido Jimaku Sakusei Henshu Shisutemu" *Nihon Onkyo Gakkai Shuki Kenkyu Happyokai Koen Ronbunshu*. 2-8-4.

Kawahara, Tatsuya. 2012. "Gikai no Kaigiroku Sakusei no tameno Onsei Ninshiki: Shugiin no Shisutemu no Gaiyo" *Joho Shori Gakkai Kenkyu Hokoku*. SLP－93－5.

Kawahara, Tatsuya. 2017. "What makes a quality transcript in parliamentary reporting" *Presented at the 51st INTERSTENO Congress,* Berlin, Germany, 22-28 July 2017.

Masuyama, Mikitaka. 2016 a. "Reducing the Space and Time between Citizens and Parliament: Video Retrieval System for Diet Deliberations" *Prepared for delivery at the 2016 General Conference of the European Consortium for Political Research*, Charles University, Prague, Czech Republic.

Masuyama, Mikitaka. 2016 b. "Linking Parliamentary Minutes and Videos in the Japanese Diet." GRIPS Discussion Papers. 16-22.

Masuyama, Mikitaka. 2017a. "Video Retrieval System for Diet Deliberations and its Experimental Applicability." (in Japanese) GRIPS Discussion Papers. 17-10.

Masuyama, Mikitaka. 2017b." Video Retrieval System for Diet Deliberations and SNS Information Sharing." (in Japanese) GRIPS Discussion Papers. 17-12.

Masuyama, Mikitaka. 2018a. "Reading and Viewing Parliamentary Debates: The Case of the Japanese Diet." (in Japanese) *Presented at the 2018 Annual Meeting of the Japanese Association of Electoral Studies*, Takushoku University, May 12-13, 2018.

Masuyama, Mikitaka. 2018b. "Reading and Viewing Parliamentary Debates : The Case of the Japanese Diet" *Prepared for the 2018 Annual Meeting of the American Political Science Association*, Boston, August 30 - September 2, 2018.

Masuyama, Mikitaka. 2018c. "Video Retrieval System for Diet Deliberations and its Experimental Application." (in Japanese) GRIPS Discussion Papers. 18-07.

Masuyama, Mikitaka and Kaori Takeda. 2014. "Instant Parliamentary Deliberations Are in Our Reach" *Presented at the 2014 Annual Meeting of the American Political Science Association*, Washington, D.C., USA, August 28-31, 2014.

Masuyama, Mikitaka and Kaori Takeda. 2015. "Ikani Mitai Kokkai Shingi Eizo ni Totatsu suruka?" *Revaiasan*. 56: 54-79.

Matsuda, Kenjiro. 2016. "Taisho Showa Senzenki no SPban Enzetsu Rekodo niokeru 'Baai' no Yomi nitsuite" *Kokuritsu Kokugo Kenkyusho Ronshu*. 11: 63-81.

Appendix: Misrecognized "Self Defense Force" ("*Jieitai*") in the House of Representatives meeting during the 197th Diet session

Meeting	Plenary 2018/10/29
Speaker	Prime Minister Abe
Match Minute Sound	今後三年かけて、子供各県子どもから若者、子育て世代、混線しない現役世代、高齢者まで、全にすべての世代が安心できる社会保障制度へと改革を進めてまいります。自衛隊と憲法改正についてお尋ねがありました。http://gclip1.grips.ac.jp/video/video/7584?t=2h41m29s
Match Minute Sound	それにもかかわらず、近年賃金における調査でも、自衛隊は合憲共済まあ自衛隊の貢献と言い切る憲法学者は二割三案にとどまり、多くの教科書に合憲性に議論がある旨の記述があります。http://gclip1.grips.ac.jp/video/video/7584?t=2h42m31s
Meeting	Plenary 2018/10/30
Speaker	Prime Minister Abe
Match Minute Sound	全ての自衛隊員は、すべての国に大変は事に臨んでは危険を顧みず、県を初め身をもって責務の完遂に務め、も監視に努め持って国民の負託に応こたえると専務宣誓し、任務二につきます。http://gclip1.grips.ac.jp/video/video/7585?t=2h4m1s
Meeting	Budget Committee 2018/11/1
Speaker	Honda
Match Minute Sound	特に陸上自衛隊の一般の隊員の皆さんの状況が非常に、どういうことなんだということがたくさんあるので、んでちょっと細かい例を、大臣、お聞かせいただきたいんですけれども、。自衛官の方が勉強する、いろいろんな作戦などを勉強する教標本というのは、自衛官の方はちゃんと自衛隊からもらえるんですか。http://gclip1.grips.ac.jp/video/video/7589?t=8h2m23s
Speaker	Minister Iwaya
Match Minute Sound	そのような話は、実は私も防衛大臣に就任する前に、地元の自衛隊全体家族会の皆さんから聞かされたこと事がございました。http://gclip1.grips.ac.jp/video/video/7589?t=8h8m3s
Match Minute Sound	十一月十四八日に開催した平成三十年度自衛隊絶対記念日観閲式に御出席いただいた国会議員につきましては、政党別に申し上げますと、。自由民主党五名、米公明党四名、立憲民主党三名、計経営十二名でございましたす。http://gclip1.grips.ac.jp/video/video/7589?t=8h9m47s
Meeting	Cabinet Committee 2018/11/22
Speaker	Shinohara
Match Minute Sound	自衛隊は、さっきあって自衛隊も先ほどありましたけれども、。専守防衛への配慮から、攻撃広域を前提とした本格的な能力の保有はないですから、。ただ専守防衛の趣旨からすれば、ほかの手段がに軍なくてどうしても攻撃保育が必要な場合には、安保体制の建前から、これは米軍に協力を求めるのかということが次の議論になってくると思うもんです。http://gclip1.grips.ac.jp/video/video/7682?t=2h6m3s
Meeting	Cabinet Committee 2018/12/5
Speaker	Shiokawa
Match Minute Sound	入間基地の拡張強化計画の撤回、安保法制の廃止を求めますが、こういった自衛隊衛生。自衛体制について、官房長官としての認識を伺います。http://gclip1.grips.ac.jp/video/video/7747?t=1h1m41s
Meeting	Security Committee 2018/11/13
Speaker	Minister Iwaya
Match Minute Sound	これからこういう事故が起こってはならないし、自衛隊まあ地方へも起こしてはならないわけですけれども、http://gclip1.grips.ac.jp/video/video/7605?t=49m29s

Match Minute **Sound**	当然ながら、全省内、全自衛隊内に周知徹底をしております。 http://gclip1.grips.ac.jp/video/video/7605?t=55m29s
Match Minute **Sound**	今度の予算では、自衛隊**絶対**施設の老朽化したものの更新に四百二十億円、そして、勤務、生活用備品等の整備費を整備し http://gclip1.grips.ac.jp/video/video/7605?t=2h31m33s
Match Minute **Sound**	南西地域には、これから地元の御理解をいただいて自衛隊の基地も順次つくらせしていただきたいというふうに思っておりますが、 http://gclip1.grips.ac.jp/video/video/7605?t=2h38m56s
Speaker	Maehara
Match Minute **Sound**	岩屋大臣、防衛省・自衛隊にこういう法律のをもとにのいわゆる中国のZTEとかファーウェイ、というとまあ米使わせるということを想定できますか、防衛省・自衛隊で。。**防衛省や自衛隊大臣まあ現在** http://gclip1.grips.ac.jp/video/video/7605?t=2h0m22s
Speaker	Wada
Match Minute **Sound**	海上自衛隊における輸送艦、又は航空自衛隊**またはその国自衛体**における輸送機、増強が必要だと思いますけれども、この。点についてお伺いしたいのがまず一点目でございます。 http://gclip1.grips.ac.jp/video/video/7605?t=2h29m49s
Meeting	Security Committee 2018/11/16
Speaker	Minister Iwaya
Match Minute **Sound**	国民の生命財産を守るべき任務を担う防衛省・自衛隊**渋滞**がこのような事故を起こすことはあってはならないことでありまして、このような事故が自衛隊の運用、訓練等に関して国民の皆様に不安を与え、防衛省・自衛隊**全体**に対する国民の信頼を損なわし**占**めるものであることを隊員一人一人が重く受けとめるべきであると考えております。今後、各種事故の防止に全力で取り組み、防衛省・自衛隊に対する国民の皆様の信頼回復に努めてまいる所存であります。 http://gclip1.grips.ac.jp/video/video/7637?t=22m5s
Match Minute **Sound**	饗庭野演習場は、防衛省・自衛隊**運用上は防衛省全体**にとっても極めて重要な演習場でございまして、長らく、滋賀県、高島市を始め、**初め**関係自治体の皆さんにお世話になってまい**汚染参**りました。 http://gclip1.grips.ac.jp/video/video/7637?t=28m57s
Match Minute **Sound**	少子化それから人口減少の中でいかに自衛隊の**十たら**人的基盤を維持していくかというのは、極めて重要な課題であると我々も思っております。 http://gclip1.grips.ac.jp/video/video/7637?t=45m17s
Match Minute **Sound**	しかし、自衛隊**世帯**は特殊な任務に当たりますので、そういう特殊な任務に従事する自衛官については独自の手当を設けておるところでございまして、これは、さらに充実をさせしていかなければいけないというふうに思っております。 http://gclip1.grips.ac.jp/video/video/7637?t=2h24m14s
Meeting	Security Committee 2018/11/29
Speaker	Honda
Match Minute **Sound**	先日、私の地元である北海道で大きな停電が起こりまして、我々民間も大変な思いをしたんですが、まさか自衛隊の基地**ました。まさか二兵隊の父**がブラックアウトしていたというのは、私もびっくりをいたしました。 http://gclip1.grips.ac.jp/video/video/7714?t=2h1m47s
Speaker	Minister Iwaya
Match Minute	現在、陸上自衛隊、**苦情を大勢単位**全部で百五十九施設ある中で、三十**六**四の施設がまだ自家発電機を備えておりません。それから航空自衛隊でも、**一施設**七十三施

Sound	設が ある中の一施設が備えておりませんので、これを早急に整備してまいりたいというふうに思っておりまして、先生御指摘のように、災害が起こった際には自衛隊の税駐屯地というのはまさに対処象の拠点になるということでございますので、電力全力でを確保するということが極めて死活的に重要だというふうに考えております。 http://gclip1.grips.ac.jp/video/video/7714?t=2h2m26s
Speaker	Shimoji
Match Minute Sound	それで、二つ目ですけれども、。このこれまでの防衛大綱とか自衛隊のまあ考え方というかデータの考え方てのは、日本海とか南西諸島とか、こういうふうなところを重点的にまあ配備をしていくというようなことであったわけですけれども、。しかし、今、中国が、宮古の沖と石垣、京都の大きと石垣まあ宮古島と石垣の間をの通って太平洋に出ていくというまあ余裕ようなケースが多くなってきました。 http://gclip1.grips.ac.jp/video/video/7714?t=3h32m26s