

キーワードから映像へ: 「会議の見える化」の紹介

松田謙次郎¹ 河原達也² 増山幹高³ 秋田祐哉⁴

¹神戸松蔭大学 文学部 ²京都大学 情報学研究科

³政策研究大学院大学 ⁴京都大学 経済学研究科 kenjiro@shoin.ac.jp kawahara@i.kyoto-u.ac.jp

mmasuyama@grips.ac.jp akita@econ.kyoto-u.ac.jp

概要

YouTube で発信されている全国の地方議会の審議映像を対象に、任意のキーワードで検索し、それが発話されているシーンをピンポイントで再生するシステムである「会議の見える化」(<https://localassembly-video.jp/local-assembly-y/>)を解説する。YouTube で配信されている審議映像 524 件について音声データを抽出し、音声認識によって発言と映像を同期させることで、キーワードを含む発言映像のピンポイント再生を可能にしている。「会議の見える化」の活用により、地方議会の情報が有権者によりアクセシブルなものになるだけでなく、社会科学・人文的の学問領域にも多大な貢献をなすことが期待される。

1 問題と背景

議会会議録のインターネット上での公開については、国会では国会会議録検索システムが 2001 年より稼働している。地方議会においても、厳密な現段階での統計値は見当たらないものの、2021 年段階で都道府県・市・東京特別区ではすべてが、町村では 68%がウェブ公開しているとされる[1]。審議映像についても国会では衆参両院の事務局で公開しており、地方議会においても自前のサーバや YouTube のチャンネルを通じて公開している。

ただし、国会とほとんどの地方議会では、会議録と審議映像はそれぞれ別箇のシステムとして運用されている。そのため会議録で特定単語を検索し、それが含まれた発言を探し出すことは可能でも、実際にその単語を含む発言を映像で視聴するためには、それぞれの会議を収めた動画から発言箇所を手作業で探さなければならない。数時間にわたる長時間会議や、検索語が使用されたすべての会議について発言箇所を視聴しようとする場合、途方もない手間と時間を要することになるという問題がある。

「会議の見える化」は、こうした事態を打開して検索語による視聴したい発言箇所のピンポイント再生（部分再生）を可能にするインターフェースである。将来的には、ユーザが YouTube で審議映像を公開している全自治体の審議映像を、地域横断的に容易に検索できることを目標に置いている。



図 1 国会審議映像検索システム画面

「会議の見える化」のいわば前身として、国会審議映像検索システムが開発されており (<https://gclip1.grips.ac.jp/video/>) [2][5], 2012 年 11 月から運用・公開されている。これは衆参から配信された国会審議映像と国会会議録を音声認識により同期することで、キーワード検索により対応する審議映像箇所をピンポイントで再生するシステムである (図 1)。キーワードが入力されると会議録の検索結果が審議映像 URL の一覧として表示され、URL をクリックすればキーワードが含まれる発話の冒頭から再生される。再生画面には自動作成された字幕が付与され、発言者プロフィールも表示される。使用されたパネルの文字認識済み情報、当該会議のキーワード抽出・自動要約、ワードクラウド表示、感情抽出などの機能も装備している。

さらにその国会審議映像検索システムの基盤とな

ったのが「字幕自動生成システム」[3]，および世界に先駆けて国会に導入された議会用音声認識システムである[4][5]．ここでは国会審議の音声 2000 時間と 2400 万語からなる会議録のテキストを自動的に対応付けした音声言語コーパスに基づき，音声認識モデルを機械学習させることで，認識精度 95%を誇る衆議院の会議録作成システムを可能にした．「会議の見える化」はこうした審議映像と会議録の文字情報の同期システムにより，キーワード検索による審議映像の部分再生を可能にし，視聴覚的情報の統合を目指すものである．

2 「会議の見える化」の概要

「会議の見える化」についてさらに詳しく見るために，その土台となった国会審議映像検索システムについて簡単に説明する[2][6]．国会審議映像を検索可能にするために必要な審議動画と国会会議録の同期情報は，上述の字幕自動生成システムで作成され，審議映像の検索インデックスが作られる．国会審議映像検索システムのデータベースには国会会議録のテキスト・データとこの同期情報が収められ，国会会議録のテキスト・データから全文検索用のインデックスが生成される．このインデックスには会議名や発言者名もメタデータとして付与されているので，入力したキーワードやメタデータの組み合わせに対応する発言検索が可能になる．そして審議動画と国会会議録の同期情報により，検索結果に対応するシーンが再生されることになる（図 2）．

ここで注意すべきは，システムは国会審議動画自体を保持しているわけではない点である．システムはあくまで審議映像から音声データを抽出し，字幕自動生成システムで同期情報の作成を行い，国会会議録のテキスト・データおよび動画の URL と合わせ，発言と映像の同期情報を登録している．国会審議映像検索システムの最新バージョンでは，検索語の他に開会日，院名，回次，発言者性別，年齢層，発言者会派，発言者地域といった属性による検索を実装しているほか，発言だけではなく，議員名や会議名による検索も可能になっている．

「会議の見える化」は，いわば国会審議映像検索システムの地方議会版と言えるが，地方議会が国会と大きく異なるのは，国会が衆参二院だけであるのに対し，地方議会の場合全国の地方議会が対象である点である．都道府県議会レベルでは自前のサーバから審議映像を発信しているが，市町村レベルの自

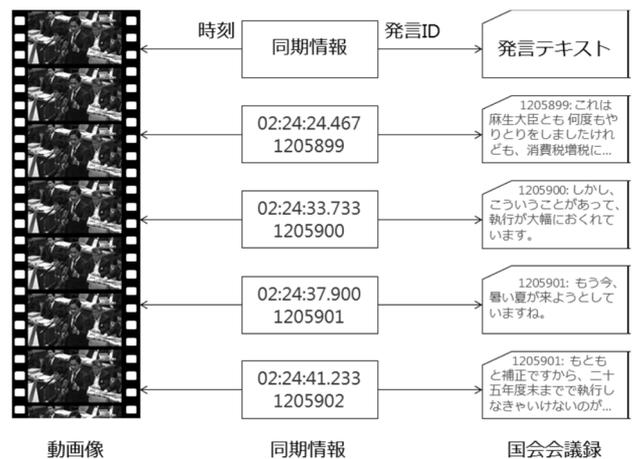


図 2 国会審議映像検索システムにおける動画と会議録の同期情報[2]

治体では，コストの問題もあつてか全体の約 1/3 が YouTube に設けられたチャンネルを通じて審議映像を配信している．そこでこのプロジェクトでは，YouTube で配信されている審議映像すべてについて，音声データを抽出し，音声認識と画像・音響解析によって発言と映像を同期させて，キーワードを含んだ発言映像がピンポイントで再生されるシステムの構築に取り組むこととした．

国会同様，ほとんどの地方議会でも会議録と審議映像は別システムとして稼働しており，一部で結合させているところでも，かなりな手作業を要しているというのが現状である．本システムとそこで使われている技術では，その自動化を目指している．

システム構築に当たり，まずは地方議会 50 件について地域横断的言語コーパスを作成し，音声認識精度を向上させることとした．この段階ではデータ量が 50 万語未満と少なく，会議録と音声認識との一致度は 79.6%に留まった．このデータから誤認識や未知語を調べたが，人名・地名以外の未知語は「ChatGPT」「DX」など少数であり，むしろ特定自治体や話者の影響が大きかったようである．次に会議録データをクロールしデータ収集を行った．

地方議会となると，方言が音声認識に及ぼす影響が懸念される．国会でも方言使用は確認されているが[7]，地方議会となると国会以上に発言中に方言が含まれることが予想される．そこでもっとも方言色が強いものと思われる沖縄県の地方議会データ（7 億単語）で検証したところ，未知語率は 0.88%であり，人名の 20%，地名の 5%が未知語であった．人名・地名以外の未知語としては，他地方でも多い「M

4 点目は収録期間である。現時点では本システムでは YouTube 上の「1 会議」が対象単位となるが、「ぎ～みる」では 47 都道府県議会では 12 年間、東京 23 区議会では 4 年間の会議録を収めている。後者が対象議会を限定し年数を深掘りするのに対し、本システムは「広く浅く」を実現していると言える。

5 点目に付加的機能である。「ぎ～みる」は地図での描画、クロス表集計、時系列グラフを提供するが、「会議の見える化」は音声認識からワードクラウドによるキーワード表示に機能を特化させている。

両システムはこのように補完的性格を有し、ユーザはその興味関心に応じて使い分けることで、両者の長点を生かすことが出来るものと思われる。

4 開発の現状

2025 年 11 月末日現在で検索可能な会議数は 524 件である。このうち約 70% (366 件) は 1 件の会議のみが収録されているが、最大で 9 つの会議が収録されている議会が 2 つあり、30%では複数会議の検索が可能になっている (表 1)。

表 1 会議の見える化: 議会毎の収録会議数分布

会議数	議会数	%
1	366	69.85
2	92	17.56
3	44	8.40
4	13	2.48
5	6	1.15
7	1	0.19
9	2	0.38
総計	524	100

524 件の会議の地域分布をプロットしたのが図 5 である。ただし地域ブロックにより都道府県数も異なるので、むしろ都道府県別に特別区・市町村総数のうちどれほどをカバーしているのかを示した方が意味があろう (図 6)。第 1 四分位以下と第 3 四分位以上の議会にラベルを付した。トップの石川県 (74%) と最下位の埼玉県 (3%) ではカバー率に 70 ポイントもの開きがあり、今後平準化を目指したい。

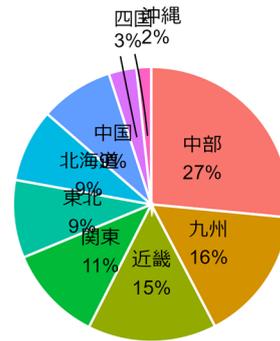


図 5 会議の見える化: 収録議会の地域分布

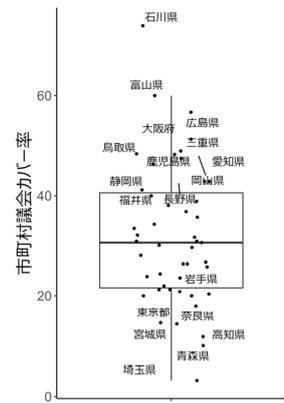


図 6 会議の見える化: 都道府県別カバー率分布

5 応用・展望

今後の開発目標の一つに審議シーンの自動識別がある。審議では議員による質疑, 担当者による答弁, 委員長発言といった多様なシーンがある。これらを自動識別し, シーンによる検索を実現させるのである。また, 発言者のプロフィールや SNS, またその発言に関連したニュースとの紐づけも可能にしたい。

この世界に類を見ない「会議の見える化」の応用として, まずはこのシステムにより, 有権者にとって地方議会の正確で効率的な情報共有が可能になり, 地方議会がよりアクセシブルになる点が挙げられる。

しかし全国に散らばる地方議会の特性, さらに映像検索という特色を活かした研究は学問領域でも活用可能である。それは直接的な関連性のある政治学や行政学に留まらず, 言語学や方言学でも強力なツールになる。発話者の声のトーン, 速さ, 大きさといったパラ言語情報は豊富な発話者の意図・感情を伝達しており, 米国議会では議員の感情と発言のピッチとの相関[10]や, *Iraq* の発音による党派性も指摘されている[11]。国会では国会審議映像検索システムを活用して「金大中」と「金日成」の議員による発音分布が検討されているが[12], 同様な調査は地方議会でも可能なはずである。審議映像があれば, さらに身振り, 手振りといったノンバーバルな伝達手段の分析にも道が開かれるであろう。

「会議の見える化」は, 議会情報を有権者に向けてよりアクセシブルなものにするだけでなく, 人文・社会科学分野にまたがる広大な学問領域において新たな知の発見を促し, 貢献をもたらすことが期待されるのである。

謝辞

本稿の作成に当たっては、内田ゆず、乙武北斗両先生より草稿にコメントを頂いた。また本研究は JST A-STEP JPMJTR23RD, および JSPS 科研費 25K04045, 22K00582 の支援を受けている。あわせて感謝したい。

参考文献

- [1] 高丸圭一, 乙武北斗, 内田ゆず, 木村泰知. 説明可能な AI 技術を用いた国会会議録と地方議会会議録の特徴比較. 宇都宮共和大学シティライフ学研究, Vol. 25, pp. 75-87, 2024.
- [2] 増山幹高・竹田香織. いかに見たい国会審議映像に到達するか?—国会審議映像検索システムの概要—. レヴァイアサン, Vol. 56, pp. 54-79, 2015.
- [3] 秋田祐哉, 河原達也. 音声認識を用いたオンライン自動字幕作成・編集システム. 日本音響学会春季研究発表会講演論文集, 2-8-4, 2013.
- [4] 秋田祐哉, 三村正人, 河原達也. 会議録作成支援のための国会審議の音声認識システム. 電子情報通信学会論文誌, Vol. J93-D, No.9, pp. 1736-1744, 2010.
- [5] 河原達也. 議会の会議録作成のための音声認識—衆議院のシステムの概要—. 情報処理学会研究報告, SLP-93-5, 2012.
- [6] 増山幹高, 国会審議映像検索システムとその実験的応用の試み. GRIPS Discussion Paper, Vol. 18, No. 7, 2018.
- [7] 松田謙次郎, 「方言」の飛び交う国会審議. 愛知大学人文社会学研究所 2017 年度ワークショップ報告書, p. 77-94. 2019.
- [8] 高丸圭一, 乙武北斗. 地方議会会議録における探索的方言調査の試み—機械学習と XAI 技術による方言研究—. 方言の研究, Vol. 9, pp. 27-51, 2023.
- [9] Yasutomo Kimura, Keiichi Takamaru, Takuma Tanaka, Akio Kobayashi, Hiroki Sakaji, Yuzu Uchida, Hokuto Ototake, and Shigeru Masuyama. Creating Japanese Political Corpus from Local Assembly Minutes of 47 Prefectures. **Proceedings of the 12th Workshop on Asian Language Resources**, pp. 78-85, 2016.
- [10] Dietrich, Bryce J., Matthew Hayes, and Diana O'Brien. Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech. **American Political Science Review**. Vol. 113, No. 4, pp. 941-962, 2019.
- [11] Hall-Lew, Lauren, Elizabeth Coppock, and Rebecca L. Starr. Indexing Political Persuasion: Variation in the *Iraq* Vowels. **American Speech**. Vol. 85. pp. 91-102, 2010.
- [12] 増山幹高, 松田謙次郎. 国会審議における同形異音語の分析. GRIPS Discussion Paper, Vol. 23, No. 4, 2023.