

# メディアアクセシビリティ **特集号**

## 巻頭言

### 語られない大切なこと

早稲田大学 人間科学学術院 教授

巖淵 守

2

## 解説

### メディアアクセシビリティ支援技術の最新動向

4

## 報告

### 競技データを利用したスポーツ手話CG生成システムの開発

20

### スポーツ中継を対象とした解説音声制作・配信システム

30

### 人と一緒にテレビを視聴するロボットの発話文生成手法

36

## 研究所の動き

### コンテンツ発見技術 ～“放送”への多様なアクセスの実現を目指して～

46

### 長寿命で省電力な有機ELディスプレイを目指して

47

### 3次元空間の撮影位置自動選択技術

48

論文紹介 49

発明と考案 50

学会発表論文一覧 56

研究会・年次大会等発表一覧 58

## 語られない大切なこと

巖淵 守

早稲田大学 人間科学学術院 教授

2022年11月のChatGPTの登場以来、連日、生成AIの議論が聞かれるようになった。本特集号で議論されるアクセシビリティに関連しても、近年、障害のある人や高齢の人など、困難を抱える人に役立つ支援技術へのAIの応用も着実に進歩してきた。例えば、大規模言語モデルと連携して、スマートフォン（以下、スマホ）のカメラに写る状況を視覚障害のある人向けに詳細に説明してくれる支援アプリや、テキストからリアルな手話映像を自動生成するサービス、コミュニケーションに障害のある人のために生成AIを利用して少数のキーワードから発言を作成する会話補助システム、ウェブサイト（ホームページ）を自動解析して、画像に対する代替テキストを自動で付与するアクセシビリティ対応のサービスなど、さまざまな製品・サービスが登場してきている。本特集で紹介されている研究にもAIが活かされているように、今後も、アクセシビリティ向上に向けてAIの応用が発展すると期待される。

一方で、急速な技術発展の恩恵に取り残されてしまう人をどのように救うかがこれまでも増して重要になることを忘れてはならない。コロナ禍の中、ALS（筋萎縮性側索硬化症）\*1を患う女性から支援を求める相談を受けた。病気が進行してボタンを押すことができなくなり、楽しみにしていたテレビやビデオ、音楽を自ら選択して視聴することができなくなったとのことであった。特にコロナ禍においては、屋内で過ごす時間が増え、映像・音楽コンテンツの視聴が多くの人にとって一層重要な役割を果たしたものの、彼女は、その楽しみの自由を少しずつ失ってしまっていた。また、大きな声が出せなくなり、隣の部屋にいる家族を呼ぶことができない不安も増していた。こうした状況の中、コロナ禍によって相談できる支援者とのつながりが減り、何もかもあきらめるしかなかったと、そのつらさを教えてくれた。AIが普及する時代を迎えながら、彼女のように必要な技術を届けてくれる人が周囲に誰もいない状況に置かれている人はいまだに少なくないと改めて思い知らされた。

身の回りの家電やナースコールの操作を可能とするためのスイッチ利用についての彼女の要望に対して、多数の病院で活用されている優れた製品の1つであり、わずかな動きで操作可能なセンサー型のスイッチを導入してはどうだろうか、と私は考えた。だが、生成AIをはじめとする急速な技術発展を遂げた現在でも20年以上も前からあるこの技術が、最適な選択肢なのか、新たな選択肢も提案するなど、我々ももっと果敢に課題解決に挑むべきではないのか、と支援チームの他のメンバーからは再検討を促された。

最終的に、スマホのカメラを利用して、顔の動きをAIで捉え、スイッチ操作に変える新たなアプリを開発し、その活用を提案した。なお、似た機能としては、スマホに標準で備わるアクセシビリティ機能に含まれる、顔の動きをトラッキングして操作に利用する機能がすでに存在する。しかし、この機能は、スマホの操作用に意図されたものであり、今回の目的には必ずしも使いやすいわけではなかった。例えば、この機能を利用するために顔のすぐ前にスマホを設置する必要があるが、顔の前の位置は介助の邪魔になってしまう。そこで、開発したアプリでは、ベッドから離れた位置に設置したスマホのカメラでも顔の

1997年 大阪大学大学院基礎工学研究科博士後期課程  
修了  
1997年 University of Dundeeコンピューター応用  
学科 助手  
2001年 広島大学大学院教育学研究科 専任講師  
2003年 University of Washington大学院教育学研  
究科 客員(文科省在外) 研究員(～2004年)  
2006年 広島大学大学院教育学研究科 助教授  
2008年 東京大学先端科学技術研究センター 特任准教  
授  
2009年 東京大学先端科学技術研究センター 准教授  
2018年 早稲田大学人間科学学術院 教授  
現在に至る

AI・DX時代における福祉とICT利用について考え、スマホやクラウドサービスなど、身の回りにあるテクノロジーを利用した支援技術の研究開発に従事。インクルーシブ社会の推進を目指し、障害のある人をはじめ、さまざまな困難を抱える人々の多様なコミュニケーションや学び方・働き方を支援する技術の開発・利用に取り組む。著書に「黙って観るコミュニケーション 重度・重複障害の子ども達とのコミュニケーションのポイント」(株式会社atacLab、2016年)、「バリアフリー・コンフリクト」(東京大学出版会、2012年)など。博士(工学)。



動きを捉えて機能するように設計した。わずかな違いに過ぎないが、現場においてはこの差が効いてくる。ナースコールとしての呼び出しやテレビのチャンネル切り替えなど、よりシンプルな目的・仕様で十分なのだが、簡単な機能ほどかえって注目されず、提供もされていない。今後、AI技術のポテンシャルを活かすために、その用途について、効率化や正確性とは異なる議論が一層重要になりうる場合は多々あると考える。彼女の例のように、語られていないニーズがいまだ多くあることに、開発者は改めて気付くことが求められるのでは、と自戒を込めて考えたい。

AIの応用で大切にしたい点について、支援技術に関する情報発信を共に行ってきた当事者である若者からのメッセージにヒントがあるように感じる機会があった。この若者は、SMA(脊髄性筋萎縮症)\*<sup>2</sup>という、先に紹介した人と異なる病気を抱えているが、全身の筋力低下を伴う点など、共通する困難がある。この若者は、筋力低下が進む中で、パソコンなどの機器をどうすればもっと楽に使うことができるかについて、幼い頃から身近な作業療法士に相談してきた。幸運にも、その人は我が国の支援技術分野で最も著名な専門家の一人であったが、回答はいつも決まって「あなたは どう思う？」から始まったのだそうである。ChatGPTに同じ質問をして、この回答が示されれば、非難的となるであろう。この若者<sup>いわ</sup>曰く、「私を含む誰よりも詳しいあなた(専門家)が、なぜ直接答を教えてくださいませんか?!」と、子どもの頃は腹立たしくて仕方なかったそうである。しかし、その返答の意味を、今なら分かるのだそうだ。この若者の場合、あの導きがあったからこそ、自分にとって大切なこと・道具・工夫を自ら調べ、考えることとなり、その繰り返しによって、他の人にも支援技術を紹介できる今の自分になれたのだと。また、単に模範解答とされる技術を提供されることが、必ずしも最善の答えになるとは限らないことを、自身の経験も踏まえて語ってくれた。ニーズは、人によってさまざまである。問いを投げかけたとき、最も確からしい答えを持っていながらも「あなたは どう思う？」と、今後AIが切り出してくれることもあるだろうか?

\*1 脳と脊髄にある運動神経細胞の異常により全身の筋力が低下していく病気。

\*2 脊髄にある運動神経の遺伝子の変性により筋肉が弱っていく病気。

# メディアアクセシビリティ 支援技術の最新動向

金子浩之      高橋正樹      奥田 誠

視覚や聴覚に障害のある方や高齢者を含む、あらゆる人がメディアの提供する情報を享受して、コンテンツをより楽しく視聴できるように、情報アクセシビリティの向上が求められている。本稿では、メディアアクセシビリティ支援技術と関連する研究開発動向について解説する。

## 1. はじめに

高齢者や障害者を含む、あらゆる人々が積極的に参加・貢献できる共生社会の実現に向けて、必要な情報が平等に得られ、意思疎通を円滑にしていく情報のバリアフリー化が求められている。

2006年12月に国連総会において障害のある人の人権や自由を守ることを定めた「障害者の権利に関する条約」、いわゆる「障害者権利条約」が採択されて以降、日本は2014年1月に批准し、さまざまな国内法制度の改革や整備が進められている<sup>1)</sup>。2015年には、国連は2030年までに持続可能でよりよい世界を目指す国際目標として、「持続可能な開発目標（SDGs：Sustainable Development Goals）」を定め、地球上の「誰一人取り残さない（Leave No One Behind）」という理念を掲げている<sup>2)</sup>。2016年4月に施行された「障害者差別解消法」や「誰一人取り残さない」を理念とするSDGsなどの社会的背景もあり、メディアにおいても障害者が必要とする情報を十分に取得できるアクセシビリティへの合理的配慮が求められている。

また、日本では少子高齢化や核家族化などが進んだことにより、65歳以上のひとり暮らしの人口が増加傾向にある<sup>3)</sup>。超高齢化社会において、正確で豊かな情報を十分に取得して生活の質を高め、社会的な孤立を防ぐためには、多様な情報へのアクセスや内容理解を支援し、話し相手のいる環境を実現していくことが必要となる。

近年、スマートフォンの普及や通信技術の高度化により多様なデジタルサービスが進展している一方で、緊急時に最新の情報を取得したり、同時に複数の方と楽しさや感動を共有するなど、テレビは重要な情報源として活用されている。NHKは、公共メディアとして、確かな情報・サービスを一人ひとりに届け、分断化・多層化した社会をつなぐ役割を果たしていく必要がある。

テレビを含むメディアには、友人や家族とコンテンツをともに視聴して楽しさや気づきを共有するコミュニケーションツールとしての側面があるが、インターネット技術の進化やモバイル端末の普及に伴い、友人や家族とコンテンツを一緒に視聴する機会が減少している。多様な情報へのアクセスや内容理解を支援したり、音声対話を利用してコンテンツをより楽しく視聴するためのパートナーとして、人と一緒にコンテンツを視聴

するコミュニケーションAIデバイスの開発が進められている。

そこで本稿では、視聴覚障害者に対するアクセシビリティ支援技術と、コンテンツ視聴時に楽しさや気づきを与えるコミュニケーションの提供を目指した技術の最新動向について紹介する。

## 2. 聴覚障害者向けアクセシビリティ支援技術の研究動向と手話CG

### 2.1 聴覚障害者を取り巻く状況

厚生労働省が2006年に実施した身体障害児・者実態調査結果によると、日本国内の聴覚・言語障害者の数は約34万人とされている<sup>4)</sup>。このうち、手話を日常のコミュニケーション手段としている方は約6万人と言われている。聴覚障害者向けの情報アクセシビリティ支援において、音声で提供される情報を補う方法には「字幕」と「手話」がある。手話は日本語と文法体系が異なる自然言語である。先天的に聴覚のない方や幼い頃に聴力を失った方にとっては、手話が母語（第一言語）であり、字幕として提供される日本語は第二言語になる。このため、手話を母語とする方から、日本語の字幕に加えて手話による情報提供が求められている。

### 2.2 字幕・手話放送の現状と課題

総務省は、「放送分野における情報アクセシビリティに関する指針」（2018年2月策定、2023年10月改定）（以下、行政指針）において字幕放送、解説放送および手話放送の普及目標を定め、放送事業者に対して普及に向けた取り組みを促している<sup>5)</sup>。2023年10月に改訂された行政指針の普及目標では、字幕放送はNHK（総合）について字幕付与可能な放送番組を対象にすべてに字幕付与すること、NHK（Eテレ）についてできる限り目標に近づくよう字幕付与することを目標としている。一方、手話放送では、2027年度までにNHK（総合）について1週間当たり平均30分以上、NHK（Eテレ）について平均5時間以上に手話を付与することを目標としている。

現在の普及状況として令和4年度の実績は、字幕放送では普及目標の対象となる放送番組における字幕番組の割合は、NHK総合で100.0%、Eテレで95.4%（総放送時間に占める字幕番組の割合は、NHK総合で90.6%、Eテレで85.9%）であり、字幕による情報保障は着実に進められている。一方、手話放送はNHK総合で1週間あたり59分、Eテレで4時間2分にとどまっている<sup>6)</sup>。このように字幕放送と比べると手話放送の割合は低いことから、手話サービスのさらなる拡充が求められている。

放送番組への手話付与では、ニュースや災害、スポーツなどの多様な話題について、手話表現への高い専門性や技能が必要とされる。現在、さまざまな機関・団体で手話通訳の育成が進められているが、手話通訳の人数は限られ、手話付与には体制面や運用面で多くの課題がある。放送局に手話通訳を常時確保することは難しく、特に災害時に緊急で人員を手配し、手話付与の体制を構築することは極めて困難である。このため、地震や津波などの緊急時でも人手をかけることなく、字幕に加えて手話による情報提供を行える技術的手段が求められている。

1表 EASIER参加機関

参加機関	国	担当
マーテルイノベート社 (Martel Innovate)	スイス	監督、進捗管理
アテナ研究センター (Athena Research Center)	ギリシャ	アバターアニメーション生成、ギリシャ手話評価実験
ハンブルク大学 (University of Hamburg)	ドイツ	対訳コーパス提供、ドイツ手話評価実験
ラドバウド大学 (Radboud University)	オランダ	オランダ手話
サリー大学 (University of Surrey)	イギリス	手話認識・アノテーション
チューリッヒ大学 (University of Zurich)	スイス	音声言語-手話翻訳
フランス国立科学研究センター (French National Centre for Scientific Research)	フランス	CG生成ポスト処理
ドイツ人工知能研究センター (German Research Center for Artificial Intelligence)	ドイツ	CG生成ポスト処理
オーディアリング社 (audEERING)	ドイツ	テキスト・音声の感情分析
ニューロメディア社 (Nuromedia)	ドイツ	手話翻訳アプリ開発
スイステキスト社 (SWISS TXT)	スイス	手話放送映像提供、コマーシャライズ
欧州ろう連合 (European Union of the Deaf)	ベルギー	評価実験
インタープレティス社 (Interpretis)	フランス	フランス手話
ユニバーシティ・カレッジ・ロンドン (University College London)	イギリス	対訳コーパス提供、イギリス手話評価実験

### 2.3 手話CG生成技術の研究動向

手話による放送・サービスの拡充に向けて、音声言語をもとに手話映像を生成する技術の研究開発がさまざまな機関で進められている。この分野について継続的に取り組まれている研究開発事例について動向を解説する。

#### (1) 多言語の手話翻訳研究に関する欧州プロジェクト (EASIER)<sup>7)</sup>

EASIER (Intelligent Automatic Sign Language Translation) は、1表に示す14の研究機関が参加して、2021年1月から2023年12月まで行われた欧州連合 (EU) の研究資金助成プログラムHorizon 2020のプロジェクトである。このプロジェクトでは、多言語のテキストから多言語の手話CGを生成する多言語手話翻訳システムを構築することを目標として、翻訳・CG生成・アノテーション\*1・学習データ構築などの研究分野を分担して研究開発を進めた。手話を表現するポーズは人手の手作業により作成し、複数のポーズから動作を補間するキーフレームアニメーションで動作を生成するため、手指の動作や表情の表現力に課題がある。最終的にはイギリス手話、フランス手話、イタリア手話、ドイツ手話、スイスドイツ手話、オランダ手話、ギリシャ手話など多言語の手話翻訳を目指しているが、現在はドイツ語-ドイツ手話の限られた言語を対象に、旅行の話題について翻訳・CG生成が可能となっている。

このほか、EASIER参加機関のサリー (Surrey) 大学では、スピンアウト企業の Signapse社とともに、生成AI技術を用いた実写ベースの手話動画の翻訳・生成技術の研究開発が進められている<sup>8)</sup>。音声言語のテキストをもとに手話グロス列\*2に翻訳し、

\*1  
手話表現について記述したデータを書き起こすこと。

\*2  
手話を表現する手話単語列であって、手話単語の識別をするために付けた意味の近い音声言語のラベルで表記したもの。

さまざまな人物で構成される各手話単語の映像シーケンスを生成する。次に、手話単語の映像シーケンスを結合してフレーム選択ネットワーク (FS-NET: Frame Selection Network)\*<sup>3</sup>により冗長なフレームを削減するとともに、スケルトン (骨格) で表現される手話姿勢シーケンスを生成する。特定の人物のスタイル画像として指定することで、敵対的生成ネットワーク (GAN: Generative Adversarial Network)\*<sup>4</sup>を用いてフォトリアリスティックな手話動画を生成することが可能である<sup>9)</sup>。

\* 3  
連続した手話グロス列を表現する手話映像シーケンスを学習して、適切な映像フレームを選択すること。

\* 4  
生成と識別の2つのネットワークが相反した目的で学習する教師なし学習アルゴリズムのひとつ。

## (2) 手話翻訳アプリの開発に関する欧州プロジェクト (SignOn)<sup>10)</sup>

SignON (Sign Language Translation Mobile Application and Open Communications Framework) は、EASIERと同様に、2021年1月から2023年12月まで行われた欧州連合 (EU) の研究資金助成プログラムHorizon 2020のプロジェクトである。アイルランドのダブリンシティ大学 (Dublin City University) やオランダのティルブルフ大学 (Tilburg University) など、17の研究機関が参加している。

このプロジェクトでは、聴覚障害者と聴者のコミュニケーションのための手話翻訳アプリの開発を目的としている。アイルランド手話、イギリス手話、オランダ手話、フランス手話、スペイン手話への対応を目指している。

音声言語から手話への翻訳では、単語間の意味的關係をグラフ構造で表現する抽象的意味表現 (AMR: Abstract Meaning Representation) を用いて表現し、それを手話で表現するためにSiGML (Signing Gesture Markup Language) という手話記述言語に変換して、手話CGの動作を生成する<sup>11)</sup>。記述された手指や身体の動きにしたがってアバターを制御するため、手話CGによる表現の汎用性は高いものの、自然な人間の動きとかけ離れた機械的な動作が生成される課題がある。

## (3) 韓国、日本での手話翻訳研究動向

EQ4ALL社は、2017年に設立された韓国を拠点とする手話CG技術のベンチャー企業である。ルールベース\*<sup>5</sup>の機械翻訳エンジンを開発し、交通機関や公共の場の情報端末での手話CGによる情報提供やAndroid TVの手話語学教育サービスなどを実施している<sup>12)</sup>。2023年に開催された第19回世界ろう者会議ではVRを活用した手話CGによる美術館や観光地における情報提供サービスの事例を展示し、放送機器展覧会NAB Show 2023ではETRI (Electronics and Telecommunications Research Institute: 韓国電子通信研究院) と連携して次世代の地上放送方式であるATSC3.0 (Advanced Television Systems Committee 3.0) による緊急警報情報の開発成果を展示した。現在は韓国語から韓国手話へのニューラル機械翻訳の研究に着手し、ETRIでは学習データ拡充のための画像認識による手話グロス認識の研究を進めている。

\* 5  
あらかじめ整理した条件に従って、入力された内容を判別し、出力を生成する手法。

当所においても、聴覚障害者向けの情報アクセシビリティ支援技術として、人による手話通訳で対応できない部分をCGアニメーションにより代替し、手話付与の機会を増やすことを目的として、手話CG生成技術の研究開発を進めている<sup>13)</sup>。これまでに手話辞典などに掲載されている約8,000語の手話単語をモーションデータとして撮りためている。手話CG単語は、検索サイトとして公開し、多くの手話学習者に動画で確認できる手話単語の辞書としても利用されている<sup>14)</sup>。また、気象やスポーツの情報を対象に、あらかじめテンプレートを用意することで手話CGを自動生成する技術を開発した。天



1図 天気・防災 手話CGサイト 天気予報を伝える手話CG

気・防災手話CGサイトでは、気象庁から発表されたデータを基に手話CGを自動生成し、災害時にいち早く避難・警戒を呼びかける情報を提供している（1図）<sup>15)</sup>。また、NHK手話ニュースの手話映像に対して日本語音声と映像中に表現された手話を単語ごとに書き起こした15万文対の学習データを蓄積し、任意の日本語ニュース文をもとに翻訳して手話CGを生成する技術の研究を進めている<sup>16)</sup>。

手話を母語として活用されている当事者のなかには、CGアバターによる手話アニメーションに対して嫌悪感を示す方もいる。その一方、社会で手話による情報提供のある場面は少なく、表現力が高く、伝わりやすい手話CGであれば、受け入れられる可能性もある。ろう者の意見を取り入れながら表現力の高い適切な手話CGの研究開発を進めていくことが求められている。

### 3. 視覚障害者向けアクセシビリティ支援技術の研究動向と解説音声

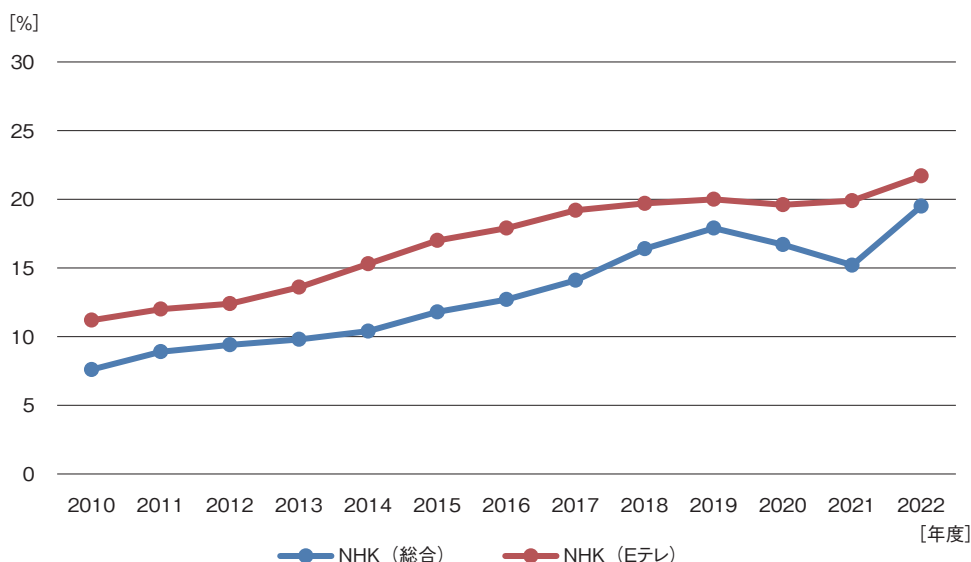
#### 3.1 視覚障害者を取り巻く状況

視覚障害とは「視機能の永続的な低下により、学習や生活に支障がある状態」を指し、ある程度の視力がある場合と失明している場合が含まれる<sup>17)18)</sup>。厚生労働省が2016年に実施した身体障害児・者実態調査結果によると、日本の視覚障害者の数は約31万人とされている<sup>19)</sup>。内訳は、全盲者が約12万人、全盲ではない視覚障害者（弱視者）が約19万人である。しかし、これらの数は視覚に関する障害者手帳の交付数を基に算出したものであり、日常生活に支障をきたしていても障害者手帳を所持していない弱視者は100万人を超えるとも言われている。さらに、現時点で十分な視力を有する晴眼者であっても老化による視力低下は避けられず、今後の高齢化社会においてはこの数がますます増えていくことが予想される。

#### 3.2 解説放送

映画やテレビ、インターネットなどの映像メディアは、パソコンや携帯電話などさまざまなデバイスの画面を通して情報を提供している。しかし、視覚に障害を持つ方々は





2図 対象の放送番組における解説付与率の推移

これら映像メディアからの情報を十分に享受できないため、視覚以外の伝達方法により情報保障を図る必要がある。テレビの映像情報を視覚障害者の方々に伝える方法の一つに「解説放送」がある。解説放送とは、出演者の表情や動作、情景描写、場面設定などの映像の内容を、副音声チャンネルで伝える放送サービスである。NHKでは1990年からサービスを開始し、ドラマや教育番組など主に収録番組で実施されている。

総務省が策定した「放送分野における情報アクセシビリティに関する指針」<sup>20)21)</sup>では、2027年度までに対象の放送番組のNHK（総合）は15%、NHK（Eテレ）は20%以上に解説を付与することを目標としている。解説放送に対応した番組は年々増加しており、2図に示すとおり解説付与率は向上している<sup>22)</sup>。さらなる番組数の増加が期待されるが、解説放送は番組が完成した後に新たに解説放送用の台本を作成してナレーションを収録する必要があるなど、制作コストの高さが解説付与率の向上を困難にしている。また、主音声の発話の合間に的確に解説することが難しいことから、生放送番組ではほとんどサービスされていない。

これらの課題に対処するため、AIによる解説放送制作の自動化や通信を併用した解説放送サービスの多様化など、新たな技術の活用が検討されている。当所においても解説放送が困難なスポーツ中継を対象とし、テレビ映像と同期した解説音声を利用者のスマートフォンへ配信するシステムを開発している<sup>23)</sup>。発話する解説文は、オペレーターによる手動入力と映像解析による自動生成により作成する。また、解説音声発話アプリに話速や情報量選択機能を設け、ユーザーの利便性を高めている。

近年では画像を入力可能な生成AI技術も提案されており、画面の内容を説明する文章の自動生成が可能になってきている<sup>24)</sup>。解析精度や速度、情報の信頼性などの課題を解決する必要があるものの、適用すればより多くの番組に解説を付与できる可能性があり、今後の技術動向が注目されている。



3図 バリアフリーマーク（音声ガイド）

### 3.3 映画における音声ガイドサービス

解説放送のように、映像メディアの視覚的な情報を音声で伝えるサービスは「音声ガイド」と呼ばれ、映画では「UDCast」<sup>25)</sup>や「HELLO！MOVIE」<sup>26)</sup>などのサービスが提供されている。これらのサービスでは、スマートフォンやタブレット端末に専用アプリをインストールしておくことで、上映中の映画に同期した音声ガイドを聴取できる。風景や登場人物の表情などの視覚的な情報をセリフの合間に補い、視覚障害者の映画鑑賞をサポートする。3図は音声ガイドのバリアフリーマークであり、このマークが付与されているコンテンツで音声ガイドサービスを受けることができる。

音声ガイドの再生には「音響同期システム」という方式を用いている<sup>27)</sup>。映画の音声解析したデータと、スマートフォンのマイクから収録した音声情報を照合し、音声ガイドを再生する。映画コンテンツに含まれる音声データを用いたサービスであるため、劇場に専用設備は不要であり、映画の途中からでもシーンに合った音声ガイドを聴取できる。また、このシステムに対応した作品は、DVDやBlu-ray、放送、配信など、さまざまなメディアにも対応可能である。

この音響同期システムは字幕表示も可能であり、シーンに合わせて字幕を表示することで、聴覚障害者が映画を楽しむことにも貢献している。障害の有無にかかわらず、誰でも劇場観賞を楽しむことができるバリアフリー上映のさらなる普及が期待される。

### 3.4 PC環境における視覚障害者支援

#### (1) スクリーンリーダー

近年では、PCやモバイル端末を活用してWebサイトから情報を得ることが一般的となった。PC環境における視覚支援技術の一つにスクリーンリーダーがある<sup>28)</sup>。スクリーンリーダーとは、コンピューターの画面を読み上げるソフトウェアであり、「音声読み上げソフト」とも呼ばれている。テキストやボタンに関する視覚情報を音声情報に変換して読み上げることで、操作・閲覧を補助するアクセシビリティ技術である。点字ディスプレイへの出力が可能なスクリーンリーダーもあり、視覚的な情報を触覚情報として伝えることもできる。

スクリーンリーダーを使用することで、視覚障害の有無にかかわらずMicrosoft WordやExcelといった一般的な事務系アプリケーションの使用が可能になる。文書の閲覧や編集などの作業を晴眼者と同環境で実現できるため、視覚障害者の就労支援にもつながっている。また、弱視者（ロービジョン）の方々はGUI（Graphical User Interface）<sup>\*6</sup>に

\*6  
画面を見て操作するインターフェイス。

2表 WCAG2.0における4つの原則とガイドライン

原則	ガイドライン
知覚可能	代替テキスト
	時間依存メディア
	適応可能
	判別可能
操作可能	キーボード操作可能
	十分な時間
	発作の防止
	ナビゲーション可能
理解可能	読みやすさ
	予測可能
	入力支援
堅牢 (Robust)	互換性

よる視覚的な操作が負担となる場合があるが、スクリーンリーダーによる音声情報を併用することにより、作業負担を軽減することもできる。

## (2) Webアクセシビリティ

Webサイトの中には、利用環境や身体的能力を考慮せずにデザインされているものも多い。そこで、障害の有無にかかわらず誰もがWebコンテンツを閲覧・利用できるようにすることを目的として、インターネットに関する技術開発と標準化を行う国際的団体W3C (World Wide Web Consortium) が主体となり、Webコンテンツのアクセシビリティに関するガイドライン WCAG (Web Contents Accessibility Guideline) の策定が進められている<sup>29)</sup>。例えば、前述のスクリーンリーダーによる音声読み上げでは画像情報を表現できないため、「写真・イラストなど画像を使用する場合は代替テキストを付与し、何が表示されているか伝わるようにする」などのガイドラインが設けられている。

WCAG1.0からWCAG3.0 (2023年時点で草案) へと段階的にアップデートされており、WCAG2.0は国際規格ISO/IEC 40500 : 2012、および日本の国内規格JIS X 8341-3:2016とも共通したアクセシビリティに関する世界的な基準となっている<sup>30)</sup>。2表にWCAG2.0で定められている知覚可能、操作可能、理解可能、堅牢の4つの原則を示す。ガイドラインのバージョンが上がるにつれて対象とする障害やデバイスの幅も拡大しており、あらゆる方々があらゆるデバイスでWebコンテンツの情報を取得できるよう、今後も検討が続けられる予定である。

## 3.5 モバイルアプリケーションによる日常生活支援

### (1) オーカムマイアイ (OrCam My Eye)

イスラエルのオーカムテクノロジーズ社 (OrCam Technologies) が開発したOrCam My Eyeは、映像を音声化する小型・軽量のカメラである<sup>31)</sup>。カメラが撮影した文字や人物の顔、紙幣、色、時計の時間などの視覚情報を耳元のスピーカーから音声で聴取できる。使い捨てライターほどの大きさであり、利用者個人のメガネに装着して使用できるため、スマートフォンのように手に持つ必要がなく、利便性に優れている。

インターネット接続不要のため、外出時にも使用できる。例えば自宅ではインターネットや新聞などのテキストを音声で読み上げ、外出時には駅や看板の案内表示、スーパーマーケットでの食品表示を音声で確認できる。また、印刷物や本を指でさし示すと、デバイスがその画像を認識して読み上げる。これらの機能は、自動車の自動運転技術に使用される画像処理技術とAIを応用して実現している。

## (2) アイナビ (Eye Navi)

Eye Naviは、(株) コンピュータサイエンス研究所が開発した、道案内と障害物検出、歩行レコーダー機能を備えた、スマートフォン用の歩行支援アプリである<sup>32)</sup>。目的地までの方向や経路、周辺施設、進路上の障害物、歩行者信号の色、点字ブロックなどを音声で通知する。

視覚障害者の歩行をサポートする機能が詰まっていることが特長であり、人・自転車・車・トラック・点字ブロック・白線・ガードレール・横断歩道など計20種類の障害物・目標物を画像解析により自動認識する。目的地までの経路の概要、交差点の有無、施設の情報などを事前に確認することも可能である。

## (3) ボイスウォッチ (Voice Watch)

(株) 電通が開発するVoice Watchは、視覚障害者も楽しめるモータースポーツ実況を生成するAIである<sup>33)</sup>。スマートフォンのカメラから得られる画像情報およびレーシングカーのラップタイムやチームの順位などのデータを活用し、現時点のレースの状況や変化が起こる兆しを検知する。また、過去のアナウンサーの実況データを学習した発話モデルを構築しており、これらを組み合わせて実況音声をリアルタイム生成する。この技術を用いることで、障害の有無にかかわらずサーキットの観客席で誰もがレースを楽しむことが可能になる。実際の大会にて、視覚障害者の方々とともに実証実験が進められている。

## 4. 人の生活を豊かにするコミュニケーションAIデバイスの研究開発動向

### 4.1 多様なコミュニケーションAIデバイス

近年の人工知能 (AI) 技術の発展により、人の生活を支援するコミュニケーションAIデバイスが家庭に浸透しつつある。音声対話が可能なスマートスピーカーは、ニュースや天気予報、交通情報など多様な情報へのアクセシビリティを向上するだけでなく、家庭用電気製品 (家電) の制御により日常生活を支援するなど、生活の質を向上させるツールとして活用されている。また、家庭用のコミュニケーションロボットでは、動物のように振る舞ったり人と対話したりするなど、人に楽しさや癒しを与えるロボットに加え、生活の情報をもとに新たな気づき・発見を提供するような人の生活体験を豊かにするコミュニケーションロボットが開発されている。さらに近年では、多様な情報へのアクセスや内容理解を支援したり、ユーザーの視聴体験を向上したりする目的で、映像コンテンツと一緒に視聴するコミュニケーションロボットの研究開発も進められている。

## 4.2 アクセシビリティを向上するスマートスピーカー

スマートスピーカーは、対話AIアシスタントを搭載し、インターネットに接続可能なスピーカーである。人との音声対話を通して、人が聞きたい音楽の再生やニュース・天気を読み上げを行う。ほかにも、家電の操作や情報検索、メール作成、スケジュール管理などさまざまな生活シーンで人を支援する。複雑な操作は必要なく、子供から高齢者まで簡単に使える特長がある。具体的にはAIアシスタント「アレクサ (Alexa)」を搭載したアマゾン・エコー (Amazon Echo)<sup>34)</sup>や「Googleアシスタント」を搭載したグーグル・ネスト・オーディオ (Google Nest Audio)<sup>35)</sup>、「シリ (Siri)」を搭載したアップル・ホームポッド (Apple HomePod)<sup>36)</sup>がある。これらスマートスピーカーは、基本的には所定のタスクを遂行するためのタスク指向型対話を行うが、タスクの遂行を目的とせず、人に楽しさなどを与える雑談対話を行うこともでき、その性能も向上しつつある。

2014年にアマゾン・ドット・コム (Amazon.com) 社がAmazon Echoの販売を開始して以降、2022年にはスマートスピーカーの世界出荷台数は2.2億台に達した<sup>37)</sup>。2021年にMMD研究所 (株)が日本国内に住む20歳から69歳のスマートフォン所有者555人を対象に行った調査<sup>38)</sup>によると、「スマートスピーカーを所有したことがある」と回答した人の割合は21.6%に及んだ。「スマートスピーカーを所有したことがある」と回答した人の内、スマートスピーカーに対し「満足している」、「やや満足している」と回答した人の割合は80%に達するなど、今後もますます利用者が増加して行くことが予想される。

## 4.3 楽しさや癒しを与えるコミュニケーションロボット

愛らしい顔や手足を持つコミュニケーションロボットも家庭に浸透しつつある。ソニー (株)が開発したアイボ (aibo)<sup>39)</sup>やGROOVE X (株)のラボット (LAVOT)<sup>40)</sup>、Macrocat社のマイキャット (Maicat)<sup>41)</sup>は、犬や猫といった動物のように振る舞い、人に楽しさや癒しを与える。人と言語による対話を行わないものの、人の呼び掛けや接触に対し、瞳を変化させたり、鳴き声を発したり、手足を動かしたりして反応する。ロボットに搭載されたカメラの撮影映像を処理することで人を識別し、その行動パターンを変化させる。障害物を避けながら部屋の中を動き回ることもできるなど、多彩な機能を有している。アイボやZoec AI社の開発ロボットキキ (Kiki)<sup>42)</sup>は、ユーザーの接し方によりその性格を変化させる。この性格形成の仕組みは、心理学分野で提唱されたビッグファイブ理論<sup>43)</sup>\*7に基づき開発された。

一方、シャープ (株)が開発を進めるロボホン (RoBoHoN)<sup>44)</sup>や (株) MIXIのロミイ (Romi)<sup>45)</sup>、ユニロボット (株)のユニボ (Unibo)<sup>46)</sup>は、人と対話を行うことができる。ロボットにより実装されている機能は異なるが、それぞれ家電の制御や天気予報の読み上げ、食事メニューの提案、ヘルスケアなど、スマートスピーカーのように簡単な音声対話を通してさまざまな形で人の生活を支援してくれる。さらに、人と雑談対話したり、歌を歌ったり、感情を表情や動作で表出したりと、人に楽しさや癒しを与える。

2023年には、各社ともロボットに大規模言語モデルを用いた対話用生成AI<sup>47)</sup>を実装した。(株) MIXIが行った調査によると、ロミイとほぼ毎日会話をしているユーザーの内、4割超が発話内容の変化に気付き、その内「より中身のある返しをするようになった」と回答した人は81.1%に、「会話内容をより覚えて返すようになった」と回答した

\*7  
人間の性格を「外向性」、「誠実性」、「協調性」、「開放性」、「神経症的傾向」の5つの因子で説明しようとする心理学理論。

人が73.0%に、「より面白い返しするようになった」と回答した人は68.4%に及ぶなど対話機能の改善が確認された<sup>48)</sup>。

家庭用ロボットは生活支援や人への楽しさ・癒しを与える効果のほかに、高齢者の認知症進行を緩和する効果やうつ病を予防・改善する効果、家族の会話を広げたり、盛り上げたりする効果も期待されている<sup>44)49)</sup>。

#### 4.4 生活体験を豊かにするコミュニケーションロボット

人の生活体験を豊かにすることを目的としたコミュニケーションロボットの研究が、活発に進められている。杉山らはドライブをより楽しいものとするを目的に、車から見える風景を話題としてドライバーと雑談する対話ロボットを開発した<sup>50)51)</sup>。ドライブ中時々刻々と変化する車からの風景画像やドライバーとの対話履歴、車が位置する場所の周辺観光情報などを基にロボットの発話文が生成される。このロボットはドライバーに楽しさを与えるのみでなく、人が漫然と車を運転し事故を起こすことを防止することも期待されている。

神田らは人と一緒にテレビゲームをプレーするコミュニケーションロボットを開発した<sup>52)</sup>。ロボットはゲームの状況に応じて、「よっしゃ！」や「やった！」といった楽しさや喜びを表現する発話を行ったり、「弱いな」や「どうした？そんなもんか？」といった対戦相手をおおるような発話を行ったりする。また、より強くロボットと一緒にテレビゲームをプレーしているという感覚を人に持ってもらうため、ロボットは発話時に人の方向を見たり、感情表現動作を行ったりする。人にゲームをより楽しんでもらうため、ゲームの状況に応じてロボットのゲーム操作レベルを操作し、人とロボットの勝負状況がきつ抗するような演出が行われる。ゲームの状況はゲーム画面を画像処理することで取得される。神田らは実験により、人が1人でゲームをプレーするより、ロボットと一緒にゲームをプレーした方が、よりゲームプレーを楽しめることを確認した。

#### 4.5 映像コンテンツの視聴体験を向上するコミュニケーションロボット

木村らが3,600人を対象に行った調査<sup>53)</sup>によると、「テレビが家族との会話をはずませたり、場をなごませたりすると思うか？」という質問に対し、約7割が「そう思う」、「どちらかといえば、そう思う」と回答しており、多くの人々が親しい人とテレビを視聴しながら楽しく時を過ごした経験をもつことが推察される。しかし、単独世帯が増加する<sup>54)</sup>など社会の変化もあり、複数人でのテレビ視聴の減少が報告されている<sup>53)</sup>。

このような背景の下、人と一緒に映像コンテンツを視聴するコミュニケーションロボットの研究が活発に行われている。Nishimuraらは、テレビ視聴時に番組の盛り上がりの雰囲気の人と共有する対話ロボットを開発した<sup>55)</sup>。番組の盛り上がりは、テレビ視聴者がソーシャルメディアに発信したコメントを解析することで推定し、この推定結果に基づきロボットの発話行動や笑いの感情表現が制御される。ロボットの発話文はソーシャルメディアに発信されたコメントから選択される。実際に人とロボットと一緒にサッカーやお笑い番組を視聴し、提案手法の有効性が確認された。

保田<sup>ままだ</sup>らは、野球に対する人の認知的観戦能力を向上する目的で、人と一緒に野球を観戦する対話ロボットを開発した<sup>56)</sup>。認知的観戦能力とは、プレーやゲームに関する知識を活かし、観戦者が観戦を通して理解したり評価したりしながら、自身の内面に意味づ



4図 人とロボットのテレビ共視聴

けすることができる力のことである<sup>57)</sup>。ロボットは試合の状況に応じ、選手や各プレーの説明、各プレーに対する感情表出などを行う。インターネットで配信されたプロ野球映像を人とロボットと一緒に視聴した結果、ロボットが登場選手の名前やプレーの内容(確認)を発話するだけより、将来起こって欲しいプレーの内容(予測)や各プレーを批評(評価)する発話も行ったときの方が、人はより深く選手の動きの意味を理解して野球を観戦するようになるなど、人の認知的観戦能力の向上が確認された。

当所では、テレビ視聴をより楽しいものとすることを目的として、人と一緒にテレビを視聴するコミュニケーションロボット(テレビ視聴ロボット)の研究を進めている(4図)<sup>58)59)</sup>。人がテレビロボットに愛着を持ち、長期間使用するためには、ロボットが人の興味を理解し、さらにはロボットが自身の興味を持って発話することが重要であると考え、これまでテレビ視聴履歴に基づいた人の興味推定手法<sup>60)</sup>やロボットの興味獲得手法<sup>61)</sup>、これら興味と視聴番組の内容に基づいてロボットの発話文を生成する手法<sup>62)</sup>を提案した。

映像コンテンツの内容に合わせ適切に感情表現動作するロボットには、一緒にコンテンツ視聴している人の興奮やリラックスといった情動を強化する効果がある<sup>63)</sup>ことから、人に正確に伝わるテレビ視聴ロボットの感情表現動作の研究も行った<sup>64)</sup>。ロボットが人に背を向けてテレビを視聴しているときにも感情が正確に伝わるよう、テレビ視聴ロボットの背部に尻尾やライトを配置し、表出する感情ごとにこれらの制御を綿密に設計した。

テレビ視聴ロボットは人に楽しさを与えるのみでなく、以下の効果を人にもたらすことを期待している。

- 1人のテレビ視聴では得られなかった気づきや発見を人にもたらす。
- 人の興味や趣味を広げる。
- 近年社会問題になっている高齢者のコミュニケーション不足を解消したり、孤独感を低減したりする。
- テレビ視聴を通してロボットが話題を提供することで、家族や友人間のコミュニケーションを活性化し、人と人の絆を深める。

---

## 5. まとめ

本稿では、メディアアクセシビリティ支援技術と関連する研究開発動向について解説した。今後もあらゆる人がメディアの提供する情報を享受して、コンテンツをより楽しく視聴できるように、メディアアクセシビリティの研究開発を推進していく。



## 参考文献

- 1) 外務省：“障害者の権利に関する条約（略称：障害者権利条約）,”  
[https://www.mofa.go.jp/mofaj/gaiko/jinken/index\\_shogaisha.html](https://www.mofa.go.jp/mofaj/gaiko/jinken/index_shogaisha.html)
- 2) 国際連合：“THE 17 GOALS | Sustainable Development,”  
<https://sdgs.un.org/goals>
- 3) 内閣府：“令和5年版高齢社会白書（全体版）,”  
[https://www8.cao.go.jp/kourei/whitepaper/w-2023/zenbun/05pdf\\_index.html](https://www8.cao.go.jp/kourei/whitepaper/w-2023/zenbun/05pdf_index.html)
- 4) 厚生労働省：“平成18年身体障害児・者実態調査結果（平成18年7月1日調査）,”  
<https://www.mhlw.go.jp/toukei/saikin/hw/shintai/06/dl/01.pdf>
- 5) 総務省：“「放送分野における情報アクセシビリティに関する指針」の改定案に対する意見募集の結果及び同指針の改定（令和5年10月17日報道発表）,”  
[https://www.soumu.go.jp/menu\\_news/s-news/ryutsu\\_231017000168713.html](https://www.soumu.go.jp/menu_news/s-news/ryutsu_231017000168713.html)
- 6) 総務省：“令和4年度の字幕放送等の実績（令和5年12月22日報道発表）,”  
[https://www.soumu.go.jp/menu\\_news/s-news/01ryutsu09\\_02000331.html](https://www.soumu.go.jp/menu_news/s-news/01ryutsu09_02000331.html)
- 7) EASIER：“Intelligent Automatic Sign Language Translation,”  
<https://www.project-easier.eu/>
- 8) Signapse AI：“Breaking Barriers with our AI Sign Language Translator,”  
<https://www.signapse.ai/>
- 9) B. Saunders, N. C. Camgoz and R. Bowden：“Signing at Scale: Learning to Co-Articulate Signs for Large-Scale Photo-Realistic Sign Language Production” IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022).
- 10) SignON Project：“Sign Language Translation Mobile Application,”  
<https://signon-project.eu/>
- 11) V. Vandeghinste, D. Shterionov, M. Sisto, A. Brady, M. D. Coster, L. Leeson, J. Blat, F. Picron, M. Scipioni, A. Parikh, L. Bosch, J. O’Flaherty, J. Dambre, J. Rijckaert, B. Vanroy, V. Nogales, S. Gomez, I. Schuurman, G. Labaka, A. Núñez-Marcos, I. Murtagh, E. McGill and H. Saggion：“SignON Sign Language Translation: Progress and Challenges,” Proc. 24th Annual Conference of the European Association for Machine Translation, pp.501-502 (2023).
- 12) EQ4LL, <https://www.eq4all.co.kr/>
- 13) 内田, 森田, 金子 (浩), 佐野：“CGで実現する未来の手話サービス,” NHK技研R&D, No.190, pp.50-59 (2022)
- 14) NHK：“手話CG単語検索サイト,” <https://www.nhk.or.jp/handsign/searchJSL/>
- 15) NHK：“天気・防災手話CG,” <https://www.nhk.or.jp/handsign/>
- 16) NHK：“技研公開2023 アクセシビリティ支援技術,”  
<https://www.nhk.or.jp/strl/open2023/tenji/4/>
- 17) 厚生労働省：“身体障害者福祉法,” (2017)  
<https://www.mhlw.go.jp/file/05-Shingikai-12201000-Shakaiengokuyokushougaihokenfukushibu-Kikakuka/0000149291.pdf>
- 18) 文部科学省：“視覚障害,”  
[https://www.mext.go.jp/a\\_menu/shotou/tokubetu/mext\\_00801.html](https://www.mext.go.jp/a_menu/shotou/tokubetu/mext_00801.html)
- 19) 総務省：“生活のしづらさなどに関する調査（全国在宅障害児・者等実態調査）,” (2016)  
[https://www.mhlw.go.jp/toukei/list/seikatsu\\_chousa\\_list.html](https://www.mhlw.go.jp/toukei/list/seikatsu_chousa_list.html)
- 20) 総務省：“放送分野における情報アクセシビリティに関する指針,” (2018)  
[https://www.soumu.go.jp/main\\_content/000531258.pdf](https://www.soumu.go.jp/main_content/000531258.pdf)

- 21) 水野映子: “テレビの楽しみを広げるサービス,” Life Design Report, Watching, (2022)  
<https://www.dlri.co.jp/files/ld/224813.pdf>
- 22) 総務省: “令和4年度の字幕放送等の実績,” (2023)  
[https://www.soumu.go.jp/menu\\_news/s-news/01ryutsu09\\_02000331.html](https://www.soumu.go.jp/menu_news/s-news/01ryutsu09_02000331.html)
- 23) Y. Kuwano: “AI Application for Enhanced Accessibility of Live Sports Broadcast,” BroadcastAsia 2023 in Asia Tech x Singapore (2023).
- 24) K. Lin, F. Ahmed, L. Li, C. C. Lin, E. Azarnasab, Z. Yang, J. Wang, L. Liang, Z. Liu, Y. Lu, C. Liu and L. Wang: “MM-Vid: Advancing Video Understanding with GPT-4V (ision),” (2023).  
<https://arxiv.org/abs/2310.19773>
- 25) Palabra : “UDCast,” <https://udcast.net/>
- 26) Evixar : “HELLO! MOVIE,” <https://hellomovie.info/>
- 27) 川野 : “映画・映像・舞台芸術等に対応したクラウド型情報保障サービスの提供,” 国際福祉機器展H.C.R. (2022)
- 28) デジタル庁: “ウェブアクセシビリティ導入ガイドブック,” (2023)
- 29) The World Wide Web Consortium (W3C): “WCAG 2 Overview,”  
<https://www.w3.org/WAI/standards-guidelines/wcag/>
- 30) The World Wide Web Consortium (W3C): “Web Content Accessibility Guidelines (WCAG) 2.0,”  
<https://www.w3.org/TR/2008/REC-WCAG20-20081211/>
- 31) OrCam Technologies: “OrCam My Eye 2,” (2023)  
<https://www.orcam.com/ja-jp/>
- 32) Eye Navi, <https://eyenavi.jp/>
- 33) Voice Watch, <https://voicewatch-project.com>
- 34) [https://ja.wikipedia.org/wiki/Amazon\\_Echo](https://ja.wikipedia.org/wiki/Amazon_Echo)
- 35) Google Nest Audio, [https://store.google.com/jp/product/nest\\_audio?hl=ja](https://store.google.com/jp/product/nest_audio?hl=ja)
- 36) AppleHomePod, <https://www.apple.com/jp/homepod/>
- 37) 総務省 : “情報通信白書令和5年版 データ集,”  
<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r05/html/datashu.html>
- 38) 伊藤 : “スマートスピーカーに関する調査,” (2021)  
[https://mmdlabo.jp/investigation/detail\\_2007.html](https://mmdlabo.jp/investigation/detail_2007.html)
- 39) aibo, <https://aibo.sony.jp/>
- 40) LAVOT, <https://lovot.life/>
- 41) Maicat, <https://maicat.io/>
- 42) Kiki Robot, <http://www.zoeticai.com/>
- 43) P. Costa and R-R. McCrae: “Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI),” Psychological Assessment Resources (1992).
- 44) ロボホン, <https://robohon.com/>
- 45) ロミィ, <https://romi.ai/>
- 46) Unibo Business, <https://www.unirobot.com/unibo-business/>
- 47) <https://ja.wikipedia.org/wiki/生成的人工知能>
- 48) <https://prtimes.jp/main/html/rd/p/000000475.000025121.html>
- 49) 太田, 滝沢: “高齢者・認知症ケアへのコミュニケーションロボットの活用に関する文献レビュー,” 心理学評論, Vol.65, No.4, pp.395-413 (2021)
- 50) 杉山, 古賀, 西島: “移動体から見える風景を話題とする雑談対話システム,” 人工知能学会全国大会論文集, 2N5-OS-7a-04 (2022)
- 51) NTTオープンハウス2022: “移りゆく車窓の風景についてロボットと話そう,” (2022)  
[https://www.kecl.ntt.co.jp/openhouse/2022/exhibition\\_11.html](https://www.kecl.ntt.co.jp/openhouse/2022/exhibition_11.html)



かねこ ひろゆき  
金子 浩之

2002年入局。仙台放送局を経て、2006年より放送技術研究所にて、コンテンツ記述・制作および手話CG生成技術の研究開発に従事。現在、スマートプロダクション研究部シニアリード。

- 52) 神田：“テレビゲームを共にプレイして楽しませる対話ロボット,” 奈良先端科学技術大学院大学修士論文 (2023)
- 53) 木村, 関根, 行木：“テレビ視聴とメディア利用の現在 ～「日本人とテレビ・2015」調査から～,” 放送研究と調査, 2015年8月号, (2015)
- 54) 厚生労働省：“国民生活基礎調査,” (2023)  
<https://www.mhlw.go.jp/toukei/list/20-21kekka.html>.
- 55) S. Nishimura, M. Kanbara and N. Hagita: “Atmosphere Sharing with TV Chat Agents for Increase of User’s Motivation for Conversation,” International Conference on Human-Computer Interaction, Human Aspects of IT for the Aged Population, pp.482-492 (2019).
- 56) 俣田, 片上：“野球初心者の認知的観戦を補助する対話ロボットの提案,” 人工知能学会全国大会論文集, 2B6-GS-10-05 (2022)
- 57) 齊藤, 醍醐, 出口, 高岡, 中路, 島崎, 佐野, 川崎, 長野, 柳沢：“スポーツの認知的観戦能力の解明,” 体育・スポーツ経営学研究, Vol.33, pp.1-19 (2020)
- 58) Y. Hagio, M. Kamimura, Y. Hoshi, Y. Kaneko and M. Yamamoto: “TV-Watching Robot: Toward Enriching Media Experience and Activating Human Communication,” SMPTE Motion Imaging J., pp.50-58 (2022).
- 59) 星, 奥田, 萩尾, 上村, 金子(豊), 西本：“コミュニケーションロボットの発話に向けたテレビ視聴時における人同士の対話分析,” 映情学誌, Vol.77, No.1, pp.149-156 (2023)
- 60) 奥田, 金子(豊), 星, 萩尾, 上村, 西本：“電子番組表解析によるテレビ視聴者の興味推定,” 映情学技報, Vol.46, No.13, HI2022-17, pp.33-35 (2022)
- 61) 上村, 金子(豊), 奥田, 萩尾, 大亦：“個性を感じさせるテレビ視聴ロボットの開発に向けた番組への興味変化モデルの提案,” 映情学年次大, 22B-5 (2023)
- 62) Y. Hagio, M. Okuda, M. Kamimura, Y. Kaneko and H. Ohmata: “Generating Utterances for Companion Robots Using Television Program Subtitles,” Proc. 2023 ACM Int. Conf. on Interactive Media Experiences (IMX ’23), pp.254-261 (2023).
- 63) 松元, 瀬古, 青木, 宮田, 渡部, 山田：“人の情動体験を向上させる感情表出口ロボット,” 信学論, Vol.J99-A, No.1, pp.45-55 (2016)
- 64) 上村, 金子(豊), 奥田, 大亦, 佐々木, 橋田：“一緒にテレビを視聴するコミュニケーションロボットの感情表現動作の改善,” HAIシンポジウム, P-28 (2023)



たかはし まさき  
高橋 正樹

1999年入局。山形放送局を経て、2002年より放送技術研究所にて、画像認識およびアクセシビリティ技術の研究開発に従事。現在、スマートプロダクション研究部チーフ・リード。博士（情報学）。



おくだ まこと  
奥田 誠

2000年入局。大阪放送局を経て、2003年より放送技術研究所にて、機械学習、画像処理、言語処理、ロボット工学に関する研究に従事。2015年から2021年まで（国研）情報通信研究機構に出向。現在、放送技術研究所ネットサービス基盤研究部副部長。

# 競技データを利用したスポーツ手話CG生成システムの開発

内田 翼 金子浩之

## Sign Language CG Production System using Sports Data

UCHIDA Tsubasa and KANEKO Hiroyuki

### 要 約

聴覚障害者への情報保障サービスの拡充を目的とし、スポーツ手話CG生成システムを開発した。本システムは、スポーツの試合中に配信される競技データから手話CG・字幕をリアルタイムに生成し、試合進行にあわせて競技映像とともにWebページで提示するシステムである。あらかじめ実況用の日本語テキストとそれを手話に翻訳したテンプレートを用意し、試合の主催者から配信される試合状況などが記された競技データと組み合わせることで手話CGのアニメーションと日本語字幕による実況を逐次自動生成することができる。また、競技データに含まれる反則の情報を使用したルール解説のプッシュ通知機能や、見逃し視聴の際に視聴開始位置をイベントから選択できる機能などを追加し、ユーザーの視聴体験の向上を実現した。本稿では、開発システムを用いた情報保障サービスの有効性評価と、国際的なスポーツ競技大会でのサービス実施結果について報告する。

### ABSTRACT

We developed a sports-sign-language computer graphics (CG) generation system with the aim of expanding accessibility services for deaf and hard of hearing people. This system generates sign-language CGs and Japanese subtitles in real time by using data distributed by the competition organizer during a sports match, and provides them on a web page along with game video as the match progresses. By preparing Japanese text for commentary and a template translated into sign language in advance and combining them with data on the match status, the system automatically generates commentary with sign-language CG animations and Japanese subtitles. Additionally, we improved the viewing experience for users by adding a push notification function that provides rules explanations using information on violations included in the competition data and a function that allows users to select the viewing start point from a list of the events when watching on-demand streaming video. This paper reports on an evaluation of the effectiveness of our accessibility service using the developed system and the results of the service implementation at an international sports competition.

## 1. はじめに

近年、手話言語条例を制定する自治体が増加するなど<sup>1)</sup>、手話を母語（第一言語）\*<sup>1</sup>とする聴覚障害者に対して手話による情報提供への理解が広がり、放送や会見、映画や舞台などの場面で手話による情報保障\*<sup>2</sup>を目にする機会が少しずつ増えてきている。また、Webコンテンツを対象にしたアクセシビリティのガイドラインが整備されており、その中でも手話母語者に向けた手話による情報保障について言及されている<sup>2)</sup>。しかし、厚生労働省の認定する公的資格である手話通訳士の数は全国で4,075人（2024年1月4日時点）となっており<sup>3)</sup>、放送やWebを始めとするさまざまなコンテンツに手話による情報保障を付与するためには、その数が十分ではない。

手話による情報保障の事例が少ないコンテンツとして、生放送のスポーツ中継が挙げられる。理由の1つは、日本語音声から手話に翻訳するにあたって競技ごとに専門用語やルールを把握する必要があり、各競技に特化した専門知識を有する通訳者でなければ対応できない点である。手話は日本語とは異なる文法体系を持つため、日本語に対応する手話単語をそのままの語順で表出するだけの通訳では正確な意味が伝わらず、情報保障として成立しない。もう1つの理由は、スポーツ中継が長時間におよぶ点である。手話通訳は負担が大きく、長時間続けると通訳の精度が低下するおそれがあるため、原則として15～20分程度で交代する<sup>4)</sup>。スポーツの試合のように長時間にわたる手話通訳を人手で実施するためには、専門知識を有した複数の通訳者を手配する必要があり、それらが大きな課題となっている。

そこでNHKでは、気象データを利用した定型文手話CG生成技術<sup>5)~7)</sup>とリアルタイム競技データを用いた解説音声の自動生成技術<sup>8)</sup>を組み合わせ、スポーツ中継

の映像に対して手話CGや日本語字幕による実況を自動生成するシステムの開発を進めてきた<sup>9)~13)</sup>。

本稿では、これまでNHKが開発してきたスポーツ手話CG生成システムおよび評価実験の結果と、本サービスを2021年の国際的なスポーツ大会にて「手話CG実況サービス」としてインターネットで提供した際の実施結果について報告する。

## 2. スポーツ手話CG生成システム

スポーツ手話CG生成システムは、競技中に大会主催者が配信する競技データに基づき、手話CGアニメーションと日本語字幕による実況を自動生成する。競技データを利用したスポーツ手話CGシステムの概要を1図に示す。システムは競技データを解析して実況用の手話グロス\*<sup>3</sup>列と日本語テキストを生成する発話文面生成部、手話モーションデータを読み込んで手話CGアニメーションを生成する手話CG生成部、ユーザーにコンテンツを提供するコンテンツ表示部で構成される。

自動生成した手話CGアニメーションや日本語字幕を、競技映像と合成して配信することで、PCやスマートフォンなどのWebブラウザを通してユーザーに提供することが可能となる。アバターを使った手話CGであれば、手話を母語とする聴覚障害者や競技の専門家の意見を基に事前に適切な手話表現を決定することで、競技に特化した内容を手話による実況で提供できる。また、手話による実況を長時間提供し続けることも可能であり、手話

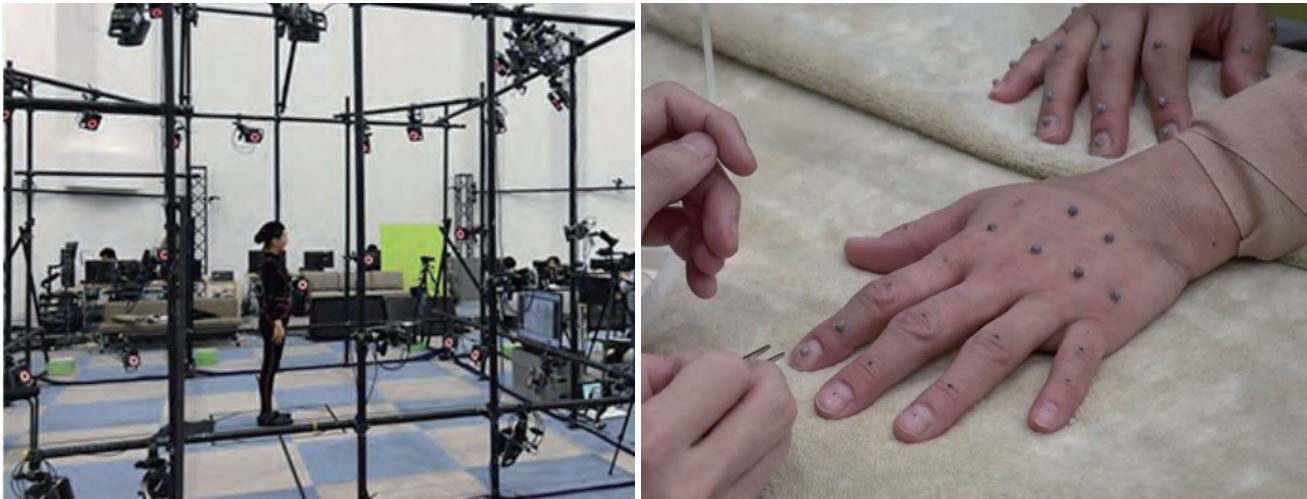
\*1 ある人が幼児期に周囲の人が話すのを聞いて自然に覚え使えるようになった最初の言語。

\*2 障害の有無にかかわらず、実質的に同等の情報が確保されるようにすること。

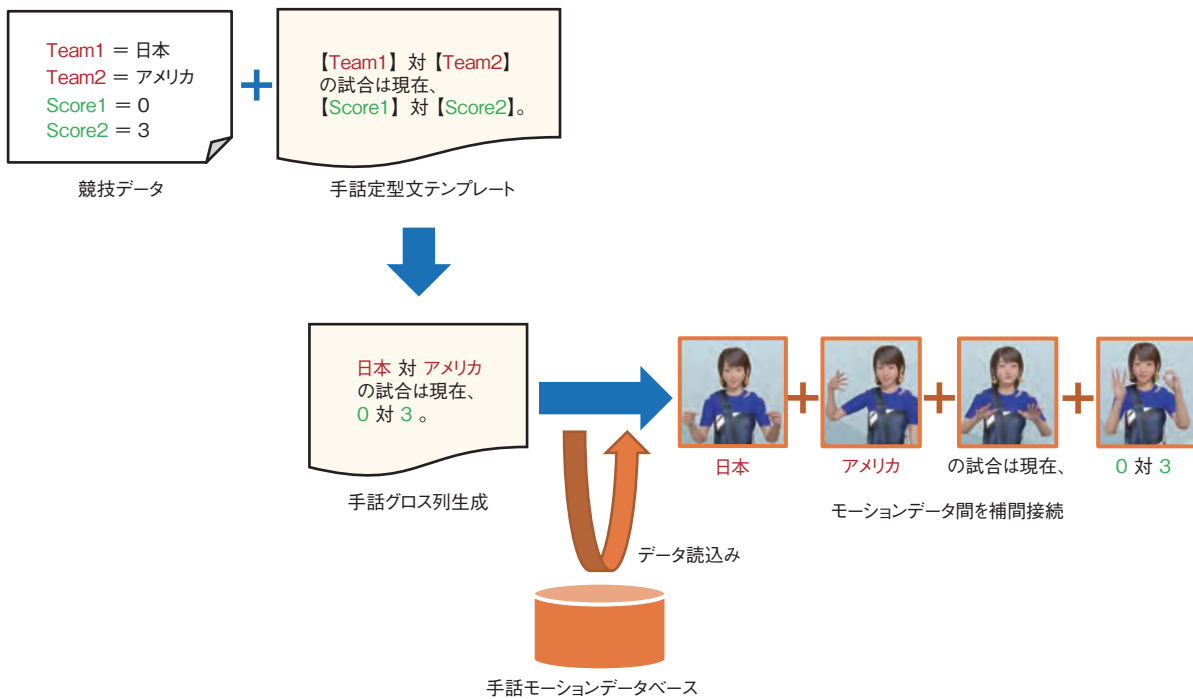
\*3 手話単語を意味的に近い音声言語（本文中では日本語）で表記したもの。



1図 競技データを利用したスポーツ手話CGシステムの概要



2図 モーションキャプチャーの様子と手に付けたマーカー



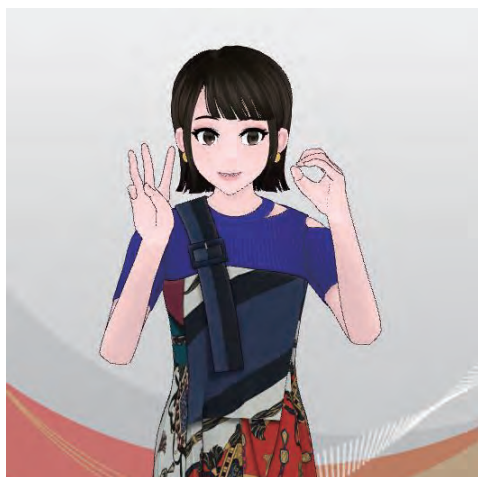
3図 手話実況発話のモーション生成フロー

付きのスポーツ中継コンテンツを自動で拡充していくこともできる。

### 2.1 手話クロス列および手話CG生成

発話文面生成部は、試合経過とともに配信される競技データを逐次解析し、対象競技ごとに事前に用意したテンプレートと組み合わせることで手話と日本語字幕用の発話文面を生成する。この仕組みではNHKが視覚障害者向けサービスのために開発したスポーツ実況音声自動生成システム<sup>8)</sup>を応用し、解説音声で使用する音声合

成用のテンプレートと同様の手話用のテンプレートを用意する。手話用の発話文面は、あらかじめ実況に用いる手話の単語や句に手話クロスを与えておき、その手話クロスをもとにテンプレートにしたがって並べた手話クロス列として記述したものである。競技データに含まれる数の表現や競技用語（例えば、ディフェンスファウル、トラベリングなど）、選手名、その他固有名詞などについても、事前に網羅的に手話クロスを与えておく。手話のテンプレート作成については、手話を母語とする聴覚障害者と競技の専門家を交えて共同で行い、もともになる日本



(a) デフォルメ



(b) フォトリアル

4図 アバターの比較

語の発話文面の意味や競技ルールをひとつずつ確認しながらユーザーにとって最適な手話表現になるように翻訳した。さらに、手話CGサービスのメインユーザーとなる聴覚障害者への事前ヒアリングにおいて、「試合に集中したいのに手話や字幕が余計な情報になってしまう」という指摘もあったため、実況として提供する発話の優先順位を検討し、不要な情報や冗長な情報を省略するといった発話内容の取舍選択も実施した。

手話CG生成部は、発話文面生成部で生成した手話クロス列に対応する手話モーションデータを読み込み、モーションデータ間を補間して接続することで1つの実況発話に相当するモーションデータを生成し、CGアバターによるアニメーションをリアルタイムレンダリングする。手話モーションデータは、手話通訳者の頭部や手足の関節などにマーカーを取り付け、光学式の手話キャプチャーシステムを用いて、顔を含めた全身の動きを収録することで事前に取得した(2図)。事前に作成した手話テンプレートをもとに、必要となる単語や定型発話フレーズのすべての手話動作をモーションデータとして用意した。手話実況発話のモーション生成フローを3図に示す。

また、CGアバターの容姿についても、事前のヒアリングなどを通じて検討した。手話では、手の形や動きだけでなく、眉の上げ下げや目の見開き、口の形(口型)、また顔全体の表情や姿勢など、さまざまな部位の動きを文法として利用しながら伝える。既存システムにおける手話CGは顔や目を大きく抽象化(デフォルメ)したアニメ調のアバターを利用していた<sup>9)~12)</sup>が、スポーツ実況サービスの実用化にあたって実写のように自然な(フォトリアル)アバターに更新した。アバターの比

較を4図に示す。フォトリアルなアバターにおいても既存のモーションデータを使って自然な表情が表出できるよう、モーションデータの変化に応じて、アバターの顔を複数の表情ターゲットを使ってブレンドするブレンドシェイプ<sup>\*4</sup>で自動調整する仕組みを取り入れた。モーションキャプチャー時の手話アクターの実写映像を参考にして、モーションデータの顔の関節の値とそのときのアバターの表情の状態を比較しながら、既存のモーションデータから最適な表情を再現するアバターの表情制御ルールを開発した。

## 2.2 コンテンツ表示

Webブラウザによるコンテンツ表示画面を5図に示す。コンテンツ表示部は、聴覚障害者への事前ヒアリングやこれまでの評価実験の結果をふまえて、手話CGによる実況だけでなく、日本語字幕や現在の対戦スコア、試合中に発生した反則のプッシュ通知やルール解説用の動画再生といった補助的な情報を提供するための機能を設けた。ルール解説用の手話CG動画は競技ごとにすべて事前に生成し、実況の手話CG動画と確実に区別できるように男性アバターを使用した。競技映像の左側には、競技映像に含まれる音声の強弱をグラフで可視化した音声メーターを配置し、会場の盛り上がりを視覚的にとらえることを可能にするなど、総合的に情報を得られるよう工夫した。また、ライブ実況後の見逃し配信用のページも設け、見逃し視聴する際に視聴開始位置を試合中に

\*4 アニメーション手法の1つであり、ボーンを使わずにメッシュの頂点を動かしてアニメーションさせる技術。表情アニメーションに多用され、あらかじめ用意した様々な表情モデルをパラメーターの調整で組み合わせることで新たな表情を生成する。



5図 Webブラウザによるコンテンツ表示画面

発生したイベントを時系列に並べたタイムラインから選択できる機能も設けた。

### 3. 評価および検証

本システムを実サービスで運用するにあたって、提供するコンテンツの有効性や生成したCGアニメーションの品質について、聴覚障害者による評価実験を実施した。さらに、評価実験の結果を受けて、改善したCGアニメーションに対する追加実験も実施した。

#### 3.1 評価実験

2021年2月18日から3月4日にかけて、Webブラウザによるコンテンツ表示の各機能と、システムで生成したCGアニメーションの品質について評価実験を実施した。

##### (1) 実験条件

実験協力者は、聴覚障害者20名（男性9名、女性11名）

で、評価はすべてビデオ通話を使った手話によるヒアリングの形式で実施した。

Webブラウザによるコンテンツ表示の各機能の評価については、サンプルコンテンツを実装した手話CG実況ページを各自操作してもらい、「音声メーター」「競技映像」「スコアエリア」「反則イベント通知エリア」「イベントタイムライン」「実況用手話CG」「ルール解説用手話CG」「実況用字幕」の8つの機能について、5段階の評価指標を選択（1：必要ない／2：許せる・分かりにくい但我慢できる／3：どちらでもない／4：改善すればもっと良くなる／5：改善する必要がない）する方法で実施した。「実況用手話CG」と「ルール解説用手話CG」の機能は、デフォルメアバターの動画で評価した。

システムで生成したCGアニメーションの品質については、手話CG動画を提示する形で2種類の評価を実施した。1つは、デフォルメとフォトリアルのCGアバターの比較評価で、同一の手話モーションデータを使って生成した2タイプのCGアバターによる手話CG動画を提示し、どちらの方が読み取りやすいか回答してもらった。



1表 コンテンツ表示に関する各機能の評価結果

機能	評点
① 音声メーター	3.40
② 競技映像	4.40
③ 対戦スコア	4.10
④ 反則イベント通知エリア	3.55
⑤ イベントタイムライン	4.35
⑥ 実況用手話CG	3.85
⑦ ルール解説用手話CG	3.90
⑧ 実況用字幕	4.20

2表 評価実験における手話CG動画の読み取り誤り率

発話文	デフォルメ (%)	フォトリアル (%)
発話文セットA	48.8	34.6
発話文セットB	51.9	53.2
全体	50.3	43.9

もう1つは、内容理解の評価で、手話CG動画を1回だけ視聴して、読み取れた選手名や国名などの固有名詞や、「シュートを決めました」や「タイムアウトを取りました」などのフレーズを手話でヒアリングする方法で実施した。競技データを本システムに入力して自動生成した実況発話30文の手話CG動画を、2タイプのアバターそれぞれで用意した。発話文を15文ごとにAとBの2セットに分けて、発話文セットA（デフォルメ）、発話文セットB（デフォルメ）、発話文セットA（フォトリアル）、発話文セットB（フォトリアル）という4つの動画提示パターンを用意し、実験協力者を5名ずつ4つのグループに分けて各グループが15動画ずつ評価した。ヒアリング終了後に、手話CG動画のアバターや手話表現に関して手話による自由回答でコメントをもらう時間も設けた。

## (2) 実験結果

Webブラウザによるコンテンツ表示の各機能の評価の結果を1表に示す。今回実装した8つのすべての機能について、実験協力者20名の平均評点が3.00以上となり、各機能の有効性が確認できた。協力者からの機能に関するコメントでは、「ルールや反則等が手話および文字で実況されるため、これまで知らなかったスポーツを知る・興味を持つ機会になる」といった好評な意見のほか、「便利な反面、一画面に掲載する情報量が多い」といった意見や「表示画面にすべて収まりきれずスクロールが必要」という意見があった。実装した各機能は必要とされているが、今後の課題として、ユーザーの好みに応じ

て各機能を表示・非表示するような選択式の提示方法が求められていることを確認した。

CGアニメーションの品質評価について、アバターの比較評価については、どちらでもよいという回答が1名で、20名中10名がフォトリアルなアバターの方がよいという結果となり、フォトリアルなアバターとデフォルメアバターの評価結果にほとんど差が見られなかった。自由回答では、フォトリアルなアバターに対する課題として、「表情がないので冷たい感じがする、親しみが感じられない」、「口型が再現できていないので別の単語に読み違えてしまう」、「指の曲がり方が再現できていないので読み取りに影響が出る」といった意見が得られた。

内容理解評価については、ヒアリングを通じて各実験協力者が読み取れなかった固有名詞やフレーズの数をもとに、発話文1文ごとの読み取り誤り率を算出することで分析した。読み取り誤り率は、発話文1文ごとに、発話文に含まれる固有名詞やフレーズの総数に対する実験協力者が読み取れなかった固有名詞やフレーズの総数の割合として定義した。各発話文において実験協力者ごとに算出した読み取り誤り率について、グループごとに実験協力者全体の平均を取った結果を2表に示す。全体として、フォトリアルなアバターの方がデフォルメアバターと比べて読み取り誤り率が6.4%少ない結果となった。しかし、発話文セットBのグループではデフォルメアバターの方が読み取り誤り率が1.4%少ない結果となり、評価に用いた発話文の内容によっては、デフォルメアバターの方が理解しやすい動作や表情表現が含まれていた可能性がある。また、自由回答において、自動生成



(a) 改善前



(b) 改善後

6図 アバターの表情制御比較

した手話CG動画について「自動生成したスコア表現の位置が低い」、「指文字表現のスピードが速い」、「単語接続時の間がないため単語がつながって見える」、「国名など初めて見る単語表現が分からない（指文字とセットで表現した方がよい）」、「表現の揺れがあるので統一した方が理解しやすくなる」といった手話表現の改善に向けた意見も得られた。

評価実験で得られた課題を受けて、アバターの表情制御機能の改善やモーションデータの調整を実施した。表情については、デフォルトでアバターの口角を少し上げて朗らかな表情を付与する工夫や、目が常にCGカメラの方向を向くような視線制御機能を追加することで、より自然な表情が再現できるようになった。6図に表情改善前後のフォトリアルアバターの表情制御の比較を示す。指摘された手話表現については、モーションデータを編集することで、口型・指の曲がり具合の改善、スコア表現における数字の位置修正、手話単語の接続点の調整や身体や頭への手のめり込みの修正などを実施した。

### 3.2 追加実験

評価実験後にアバターの表情制御機能の改善や、モーションデータを調整して生成した手話CGについて、2021年6月13日から16日にかけて追加実験を実施した。

#### (1) 実験条件

実験協力者は聴覚障害者9名(男性4名、女性5名)で、うち7名は前回の評価実験にも参加していた。評価方法は前回の実験と同じく手話によるヒアリングで、7名はビデオ通話によるオンラインで実施した。

評価項目は前回の実験同様、同じ手話モーションデータを使って生成したデフォルメ・フォトリアル2タイプのアバターの比較評価と、手話CG動画を提示し読み取れた内容をヒアリングする内容理解評価の2つを実施した。

内容理解評価で使用した手話CG動画は2種類で、システムで単語モーションを接続して自動生成した実況発話に関する女性アバターの手話CG動画5つ、単語モーションを接続せずに一文をそのままモーションキャプチャーした「バスケットボール」「車いすバスケットボール」「車いすラグビー」のルール解説に関する男性アバターの手話CG動画5つの合計10動画であった。アバターについても、デフォルメ・フォトリアルの2タイプを男女それぞれで用意した。実験協力者をアバターのタイプ別に5名(フォトリアル)と4名(デフォルメ)の2つのグループに分けて、各グループそれぞれで全10動画を評価してもらった。追加実験においても、自由回答でコメントをもらう時間を設けた。

#### (2) 実験結果

比較評価においては、どちらでも読み取りに問題ないという回答が1名、デフォルメアバターとフォトリアルなアバターがそれぞれ4名ずつという結果となった。また、男女どちらのアバターにおいても同様の結果であった。一方で、前回の実験にも参加した実験協力者からは、フォトリアルなアバターについて前回の手話CG動画よりも、口型など顔回りの表現の品質が向上したという意見が多く得られた。

内容理解評価に関しては、前回の実験同様、発話文1



7図 スポーツ手話CGシステム運用・監視中の様子

文ごとの読み取り誤り率を算出することで分析した。読み取り誤り率について、グループごとに実験協力者全体の平均を取ったところ、フォトリアルアバターの方がデフォルメアバターと比べて実況定型文で6.4%少なくなり、ルール解説文で11.1%少なくなることが確認できた。

自由回答における手話動作に関する指摘として、「定型文手話の単語結合部分や各単語の速度（特に指文字）がまだ自然ではない」というコメントがあった。本システムでは、単語や句といった複数の異なるモーションデータ間を補間して接続しているため、より自然にモーションを接続することが自然な動作につながると考えられる。また、9名中3名から「デフォルメアバターの方が輪郭がはっきりしているのを見やすい」というコメントもあった。今後の課題として、モーション接続手法の改善や、フォトリアルアバターにおける効果的な陰影表現などの必要性を確認した。

### 3.3 実サービスにおけるシステム運用

評価実験を受けて改修したスポーツ手話CG生成システムを、2021年7月から9月にかけて開催された国際的なスポーツ競技大会において運用し、NHKが公開しているWebページの特設サイト内で「手話CG実況サービス」として提供した（7図）。サービス対象は「バスケットボール」「車いすバスケットボール」「車いすラグビー」の3競技とした。競技データに含まれる情報の質や量が競技ごとに大きく異なり、実況に有用な情報の少ない競技では有意義な実況を行うことが難しいため、一定量の情報が含まれる競技を選択した。

大会期間中、すべてのコンテンツは手話通訳を一切介さずに本システムのみを用いて自動生成し、バスケット

ボールは全14試合28時間分、車いすバスケットボールと車いすラグビーの2競技は全13試合26時間分を配信した。スポーツ中継にCGの手話を付加するサービスはNHKとして初の試みであり、聴覚に障害のある方のみならず、幅広い層から好意的なコメントをいただいた。

サービス実施後に調査した聴覚障害者へのグループインタビューの結果、「言葉でペナルティーの説明があったのはよかった。言葉の意味は聴覚障害者で入ってこないから、詳しい説明があることでなるほど面白いと今回初めて思った。」（難聴女性30代）といった意見があり、サービスの意義に対して評価されていることが確認できた。しかし、「いつもは画面の下に字幕があるのに、手話CGの下にしかないから競技を見ることができない。」（難聴女性50代）といった日本語字幕の位置に関する指摘や、「動かない、動かなって気になって試合を見られない。聞きたいときに解説がない。」（難聴男性50代）といった実況のタイミングや発話内容に対する不満もあった。適切な情報の取舍選択と提示位置やタイミングについては、競技映像などの放送コンテンツと同時に手話CGを提示する際の大きな課題として、引き続き検討する必要がある。CGアニメーションについては、「手話CGの表情のうなずきとか手話の表現がすごいと思った。」といった意見が得られ、SNSでも「手話もかなり分かりやすい」、「手話に大事な微細な表情も研究され、柔らかな手の動きがすごい！」など、高品質なCGアバターによる手話表現の自然さに対して高い評価があった。一方で、「身振りでもジェスチャーでもいいのもっと表情をつけて情報を出していただきたい。」（ろう男性30代）という指摘もあり、手話を母語とする聴覚障害者にとってはアバターの表情制御についてまだ改善の余地

があると考えられるので、新たな表情制御手法を検討するなど継続課題として取り組んでいく。

## 4. まとめ

競技データを利用したスポーツ手話CG生成システムを開発し、評価実験により実装した機能や生成したコンテンツの効果を検証した。また、評価結果をふまえて機能改善したシステムを実サービスにて運用した。CGアバターをフォトリアルなアバターに更新することで、より自然で分かりやすい手話が再現可能となり、本システムで自動生成したコンテンツを使った手話CG実況サービスの意義が評価されたことが確認できた。開発したシ

ステムにより、生配信される競技映像に対して全自動で手話CGや日本語字幕による実況を付与することが可能となり、より多くの聴覚障害者に楽しんでいただけるユニバーサルなコンテンツを増やしていくことが可能となる。一方で、実況内容や提示方法、表情のさらなる高品質化に関する指摘もあり、今後本システムを利用してサービス展開する際は、ユーザーとなる聴覚障害者を交えて改善を図りたい。さらに、アナウンサーの実況を手話に翻訳して提供して欲しいといった要望もあるため、日本語から手話CGアニメーションを生成する翻訳技術の実用化に向けた翻訳精度やCG表現の品質向上に向けた研究開発にも取り組んでいく。

## 参考文献

- 1) 全日本ろうあ連盟：“手話言語条例マップ,”  
<https://www.jfd.or.jp/sgh/joreimap>
- 2) World Wide Web Consortium (W3C)：“Web Content Accessibility Guidelines (WCAG) 2.0,”  
<https://www.w3.org/TR/WCAG20/>
- 3) 社会福祉法人聴力障害者情報文化センター：“手話通訳士名簿,” <http://www.jyoubun-center.or.jp/slit/list/>
- 4) 東京手話通訳等派遣センター：“通訳者の人数について,”  
<https://www.tokyo-shuwacenter.or.jp/shuwa/haken/faq.php>
- 5) N. Hiruma, M. Azuma, T. Uchida, S. Umeda, T. Miyazaki, N. Kato and S. Inoue: “Automatic Generation System of Japanese Sign Language (JSL) with CG Animation of Fixed Pattern Weather Information,” ABU Technical Journal 264, pp.2-5 (2015).
- 6) M. Azuma, N. Hiruma, H. Sumiyoshi, T. Uchida, T. Miyazaki, S. Umeda, N. Kato and Y. Yamanouchi: “Development and Evaluation of System for Automatically Generating Sign-Language CG Animation Using Meteorological Information,” In proceedings of the 16th International Conference on Computers Helping People with Special Needs, pp.233-238 (2018).
- 7) 東, 内田：“データからの手話CG生成技術,” 映情学誌, Vol.73, No.6, pp.1040-1043 (2019)
- 8) T. Kumano, M. Ichiki, K. Kurihara, H. Kaneko, T. Komori, T. Shimizu, N. Seiyama, A. Imai, H. Sumiyoshi and T. Takagi: “Generation of Automated Sports Commentary from Live Sports Data,” In Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (2019).
- 9) T. Uchida, T. Miyazaki, M. Azuma, S. Umeda, N. Kato, H. Sumiyoshi, Y. Yamanouchi and N. Hiruma: “Sign Language Support System for Viewing Sports Programs,” In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, pp.339-340 (2017).
- 10) T. Uchida, H. Sumiyoshi, T. Miyazaki, M. Azuma, S. Umeda, N. Kato, Y. Yamanouchi and N. Hiruma: “Evaluation of the Sign Language Support System for Viewing Sports Programs,” In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, pp.361-363 (2018).
- 11) T. Uchida, H. Sumiyoshi, M. Azuma, N. Kato, S. Umeda, N. Hiruma and H. Kaneko: “Automatic Production System for Sports Program with Support Information,” In Proceedings of the 15th International Conference of the Association for the Advancement of Assistive Technology in Europe, Vol.31, No.s1, pp.134-135 (2019).
- 12) T. Uchida, H. Sumiyoshi, T. Miyazaki, M. Azuma, S. Umeda, N. Kato, N. Hiruma, H. Kaneko and Y. Yamanouchi: “Systems for Supporting Deaf People in Viewing Sports Programs by Using Sign Language Animation Synthesis,” ITE Trans. on MTA, Vol.7, No.3, pp.126-133 (2019).
- 13) 熊野, 内田, 金子：“リアルタイムデータフィードを活用したユニバーサルサービス,” 信学誌, Vol.105, No.8, pp.867-872 (2022)



うちだ つばき  
内田 翼

2011年入局。仙台放送局を経て、2014年より放送技術研究所にて、手話CGの研究開発に従事。現在、スマートプロダクション研究部に所属。



かねこ ひろゆき  
金子 浩之

2002年入局。仙台放送局を経て、2006年より放送技術研究所にて、コンテンツ記述・制作および手話CG生成技術の研究開発に従事。現在、スマートプロダクション研究部シニア・リード。

# スポーツ中継を対象とした 解説音声制作・配信システム

高橋正樹 島野雄貴 榎野裕也 宮崎 勝 佐野雅規

## Production and Distribution System for Audio Descriptions in Live Sports Broadcasts

TAKAHASHI Masaki, SHIMANO Yuki, KUWANO Yuya, MIYAZAKI Masaru and SANO Masanori

### 要 約

視覚障害者を含む多くの方々のメディアアクセシビリティ向上を目的とし、解説音声制作・配信システムの開発を進めている。本システムは解説放送サービスが困難なスポーツ中継番組を対象とし、テレビ映像と同期した解説音声をスマートフォンなどの携帯端末へリアルタイムに配信する。発話する解説文は、オペレーターによる手動入力と映像解析による自動生成により作成する。また、解説音声発話アプリに話速や情報量選択機能を設け、ユーザーの利便性を高めている。本稿では、開発システムを用いた新たな解説音声サービスの可能性や今後に向けた課題について、2022年に実施した配信実験の結果を中心に報告する。

### ABSTRACT

We have developed a production and distribution system for audio descriptions aiming for improving media accessibility for people, especially the visually impaired. The system distributes audio descriptions to users' smart phones while they are watching a live sports broadcast. The descriptive texts are created manually and with automatic image recognition. The usability of the users' application is improved by adding customization functions. In this paper, we discuss the possibility and future prospects of this new audio description service, showing the results of a distribution experiment in 2022.

## 1. はじめに

視覚障害者を含む多くの方々が映像コンテンツの内容を理解し、楽しめるよう、コンテンツに解説音声を付与する取り組みが実施されている。例えば、テレビ放送では副音声チャンネルに解説音声を提供する解説放送サービスが行われている<sup>1)</sup>。また、Webコンテンツにおいてもアクセシビリティを向上させるためのガイドラインが公開され、年々改善が図られている<sup>2)</sup>。このように、映像コンテンツへのアクセシビリティを高める情報保障サービスは近年注目を集めており、それを実現するアクセシビリティ技術<sup>3)</sup>への期待も高まっている<sup>4)</sup>。

テレビ放送における解説音声は、適切な解説内容を適切なタイミングで挿入することが難しい<sup>5)~7)</sup>。そのため、解説放送はドラマや教育番組など収録映像を主体に制作されており、スポーツ中継を含む生放送番組への適用例はほとんどない。スポーツ中継のアナウンサーは、画面上に文字スーパーされている試合情報や画面に映っている選手のプレーなど、視覚的に伝わる情報については、あえて発話しないことが多い。このような情報は視覚障害者には伝わりにくいため、テレビ音声を補足する解説音声の提供が期待されている。

そこで当所では、視覚障害者と晴眼者がテレビでスポーツ中継を一緒に視聴し、その感動を共有できるよう、テレビ映像を見ていない状況におけるスポーツ中継の状況把握を支援する解説音声制作・配信システムの開発を進めている<sup>8)~11)</sup>。中継している競技のプレーを解説

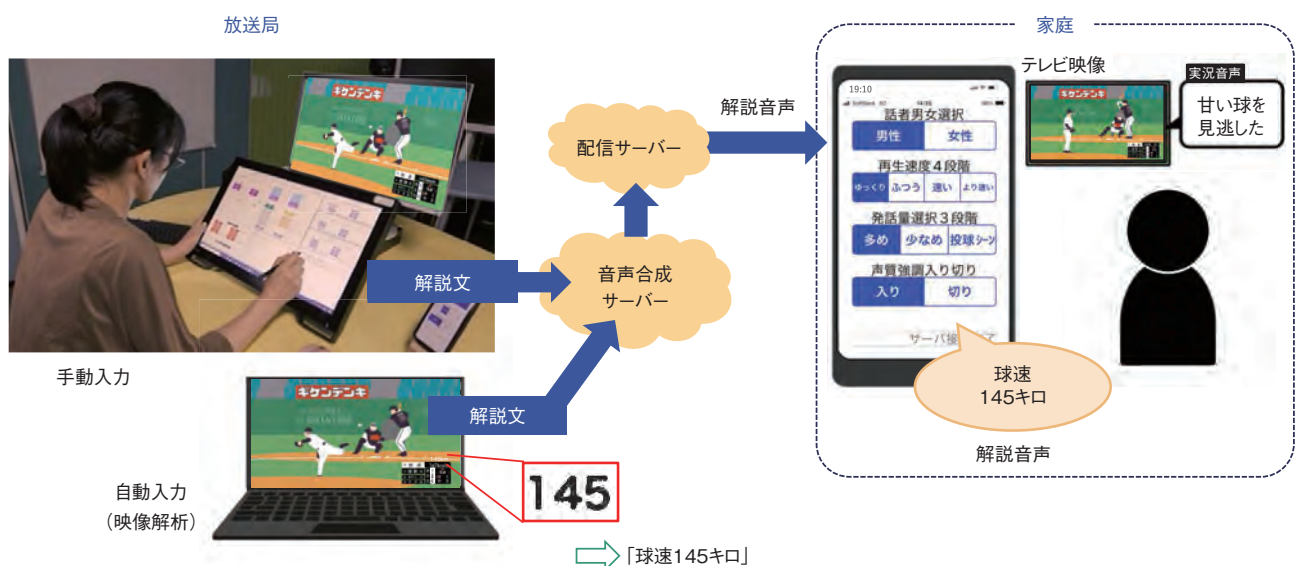
するテキストをオペレーターの操作や映像解析によって作成し、その解説文から生成した合成音をユーザーが持つスマートフォンなどの受信端末へリアルタイムで配信する。映像解析による解説文生成の自動化や解説音声発話アプリへの機能追加を通し、システムの運用性やユーザーの利便性の改善を図っている。

2021年の国際スポーツ大会、および2022～2023年のプロ野球中継にて視覚障害者を含む参加者を対象として解説音声配信実験を実施し、本サービスの有効性や追加機能の効果、今後に向けた課題等を確認した。本稿では、2022年に実施したシステム改善および配信実験を中心に報告する。

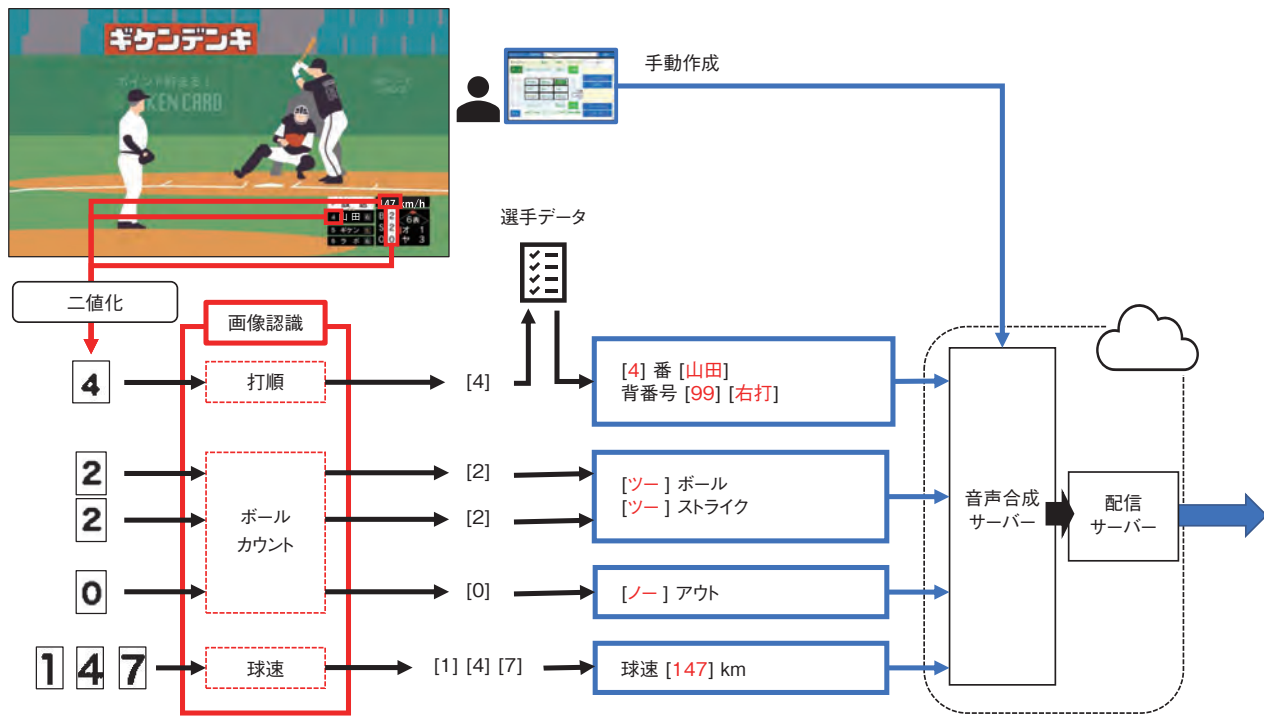
## 2. 解説音声制作・配信システム

解説音声制作・配信システムの処理の流れを1図に示す。解説音声の元となる解説文は、1図左側に示すように、人手によるデータ入力（手動入力）、および画像解析等の自動入力により作成する。解説文は音声合成サーバーにより音声データに変換され、配信サーバーを通してスマートフォンやタブレット端末などにリアルタイム配信される。家庭では、ユーザーのスマートフォンやタブレット端末に専用アプリをインストールしておくことで、テレビ映像と同期した解説音声を聴取できる。テレ

\* 1 ここでは、視覚障害者を含む多くの方々がコンテンツの内容を把握するための技術を指す。



1図 解説音声制作・配信の流れ



2図 文字認識による解説文生成の流れ

びと異なる別方向からの音源とすることで、テレビと解説音声の発話が重なる“音声被り”の影響を軽減した。さらに、専用アプリで発話者の性別や話速、声質も調節可能である。

## 2.1 解説文の自動生成機能

従来システム<sup>8)</sup>では、ほぼすべての解説文をオペレーターの操作で作成していたため、解説文の即時生成が困難であった。そこで1図の下部に示すように、映像解析により解説文の一部を自動生成する機能を実装した<sup>10)~12)</sup>。画面上で文字スーパーとして表示されている打者、ボールカウント、球速情報の文字認識、およびピッチャーの「構えた」「投げた」に関する動作認識により、解説文を自動生成した。

文字認識による解説文自動生成の流れを2図に示す。まず球速、ボールカウント、打順に関する文字スーパー表示を検知する。一般に、スポーツ中継の文字スーパーの位置やフォントは試合を通じて不変であることから、各種スーパー位置の平均輝度値をもとにスーパーのON/OFFを検知する。文字スーパーの検出後、切り出した画像に二値化処理を施し、ニューラルネットワークを用いて数字を認識する。

その後、認識結果をテンプレートの文章に挿入し、解説文を作成する。例えば、アウトカウントの認識結果が“0”だった場合、「ノーアウト」の解説文を出力する。

なお打順に関しては、認識結果と選手データを照合することで、打者に関する解説文を作成する。作成した解説文は音声合成サーバーに自動的にポストされ、音声合成を経て解説音声としてユーザーに配信される。

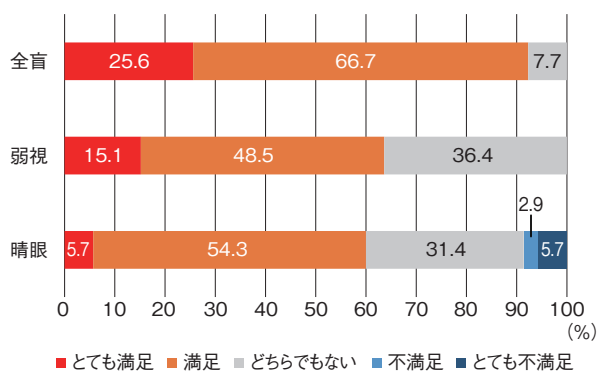
解説文の自動生成機能により手動入力負担を軽減でき、オペレーター要員を削減するとともに、オペレーターの作業に時間的余裕が生まれ、情景描写などの高度な発話をより多く提供することが可能となった。試合状況に加えて情景描写を充実させることで、解説音声サービスの向上を図った。

## 2.2 情報量選択機能

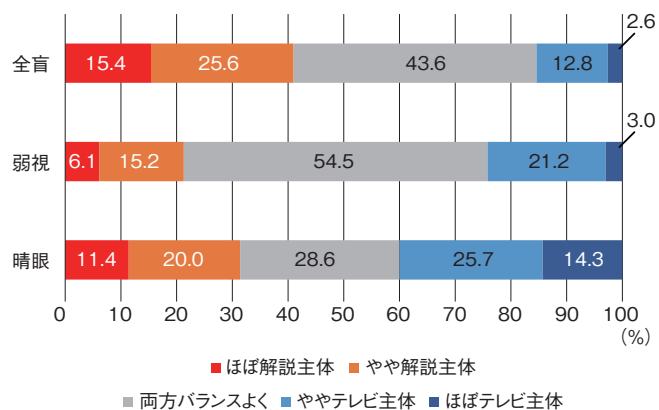
ユーザーが操作する受信端末アプリの機能として、情報量選択機能を設けた。従来のアプリでは制作した解説音声をすべて再生していたが、ユーザーによっては不要な情報もあり、カスタマイズできる機能が期待されていた。

1図右側に示すように、発話量を「多め」、「少なめ」、「打球シーン」の3段階で選択可能とすることで、ユーザーに応じて必要な情報を提供可能とした。また、情報量を調節することで、テレビ音声と解説音声をバランスよく聴取しやすくなる効果も期待した。

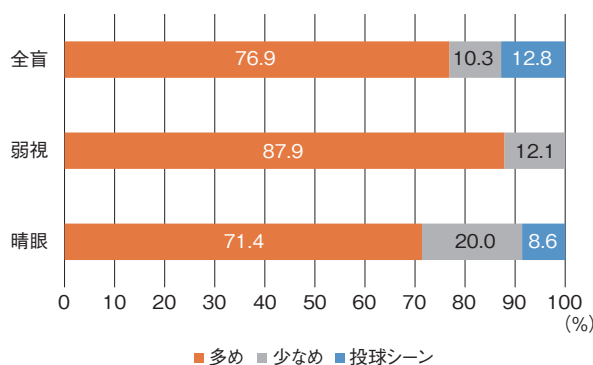




(a) 総合満足度



(b) テレビ音声と解説音声の聴取バランス



(c) 情報量選択

3図 Webアンケート結果

### 3. 配信実験

解説音声制作・配信システムの有効性や、新たに追加した解説文の自動生成機能および情報量選択機能の効果を検証するため、視覚障害者を含む参加者を対象に実際の配信を実施し、解説音声の総合評価を行った。

#### 3.1 実験条件

2022年10月に放送したプロ野球中継4試合を対象とした。全盲者12名、弱視者\*2 12名、晴眼者12名の3グループの計36名が評価者として参加し、全4試合の参加者はのべ107名となった。参加者はテレビでプロ野球中継を視聴しながら、参加者本人の受信端末から解説音声を聴取した。中継番組を視聴した後、本システムを総合的に評価するWebアンケート、およびグループインタビューを実施した。

#### 3.2 Webアンケート

Webアンケートでは、中継番組視聴後に20問の設問に対する選択式の評価と自由記述を実施し、全盲・弱視・

晴眼者のグループごとに集計した。以下に、「総合満足度」、「テレビ音声と解説音声の聴取バランス」、「情報量選択」に関する設問の評価結果を示す。

3図 (a) は「総合満足度」の評価結果である。「とても満足」から「とても不満足」までの5段階を選択肢とした。全盲・弱視・晴眼者の全体で平均すると、「とても満足」と「満足」の合計は全体の73%を占めた。特に全盲者の評価は9割以上と高く、全盲者の自由記述には、「解説音声があることで家族とスポーツ中継を楽しめた」とのコメントが複数あった。本システムは、視聴者個人に解説音声を配信することで視覚障害者と晴眼者がスポーツ中継を一緒に視聴できるよう設計しており、その目的をある程度達成できたと考えられる。

また、多様な発言内容を評価するコメントも多くみられた。「自動入力を導入したことで時間的余裕が生まれ、選手の表情や応援の様子などの情景も入力できた」とのオペレーターの見解もあり、自動入力機能が間接的に高

\*2 さまざまな病因によって眼鏡等での矯正が難しく、極度に見えづらい回復困難な視覚障害者のこと。なお本稿では、全盲と区別し弱視を医学的な分類によらず広い意味で用いている。

1表 グループインタビューの結果

	情報量	音声合成 (機械音声)	発話の遅れ
全盲	多いほどよい (聞き流せる)	抵抗なし 慣れている	気にならない
弱視	多めかつカスタマイズ 機能がほしい	抵抗なし 人間と区別できてよい	気になる
晴眼	ながら視聴では 多めがよい リクエスト機能を 付けてほしい	抵抗あり 感情を伝えてほしい	ながら視聴では 気にならない

評価に貢献したと考えられる。

3図(b)は、テレビ音声と解説音声の聴取バランスの評価結果である。テレビ音声と解説音声のどちらを主体的に聴取したか、「ほぼ解説主体」から「ほぼテレビ主体」までの5段階を選択肢とした。一方の音声を主体(ほぼテレビ音声主体/ほぼ解説音声主体)に聴取した参加者を除いた割合は全盲・弱視・晴眼者の全体平均で82.4%であり、参加者は両方の音声を聞き分けながらサービスを体験したことが分かる。テレビ音声からは過去の戦歴や解説者の感想を、解説音声からは主に現在の試合状況に関する情報を聴取し、異なる情報源として扱うことで、野球中継をより楽しめたという声もあった。

グループ別にみると、全盲者は解説音声主体の割合が比較的高く、逆に晴眼者はテレビ音声主体の割合が高い。主体とする情報が視覚障害の違いにより異なる結果となった。

3図(c)は、情報量選択の結果である。解説音声受信アプリで「多め」、「少なめ」、「投球シーン」のどのボタンを主に選んだかを選択肢とした。グループにかかわらず、「多め」を選択した参加者が大多数となった。自由記述コメントにも、より多く、より多様な情報提供を望む声が多くみられた。テレビ音声との音声被りを避ける意味でも、「少なめ」を選択する参加者が一定数いるのではないかと予想したが、より多くの情報を提供し、その上でユーザーが個々に情報量をカスタマイズできる機能の必要性を確認した。制作効率の観点からも、手動入力で作成できる解説文には限界があるため、情報量の増加に向けて解説文の自動生成を推進していきたい。

### 3.3 グループインタビュー

Webアンケートでは表れにくい参加者の生の意見を収集するため、全盲者、弱視者、晴眼者、各3名を対象とし、グループごとにリモート会議形式による事後インタビューを行った。その結果を1表に示す。

各グループで見え方が異なるため、インタビューへの

回答もグループごとに特徴が出た。音声合成の品質に関する質問に対して、晴眼者は「抑揚がなく、機械的で抵抗がある」と答えたのに対し、全盲者や弱視者は「スマートフォンでの機械音声に慣れているため、抵抗が少ない」と回答した。映像に対する解説音声のタイミングの遅れについては、弱視者はある程度画面が見えているため「遅れが気になる」と回答した。反面、晴眼者では、ながら視聴<sup>\*3</sup>でサービスを受けるケースが多かったため、「遅れは気にならない」と回答した。

一方で、情報量についてはグループにかかわらず「多いほどよい」との回答が支配的であった。ただし、ユーザーによって必要な情報は異なるため、特に弱視者でユーザーごとに情報をカスタマイズできる機能をアプリに搭載してほしいとの意見が多かった。情報量を増やしつつ、ユーザー側で情報をカスタマイズ可能な機能を実装する必要性を確認した。

## 4. まとめ

解説音声制作・配信システムに解説文の自動生成機能と情報量選択機能を追加実装し、配信実験にてそれらの効果を検証した。参加者の7割を超える方々から総合満足度で高い評価を得たことから、本サービスの有効性を確認できた。自動入力の導入により多様な情報を提供できたことが、高い評価につながったと考えられる。また、情報量をユーザー側で調整可能な機能を評価するコメントも多くみられた。一方で、より多様で、より多くの情報に対するニーズも確認できた。今後、試合イベント認識などの自動化を進め、より多くの情報を提供できるよう改善していく予定である。その上で、聴取する情報をユーザー側でカスタマイズできる機能を設け、個々のユーザーのニーズにも応えることで、アプリの利便性を高めていきたい。

**謝辞** 本研究におけるシステム開発や配信実験において、NHK財団の今井篤上級研究員および都木徹シニアエンジニアに多大なご助言をいただきました。ここに感謝の意を表します。

\*3 何か別な作業をしながら、テレビやラジオを視聴すること。

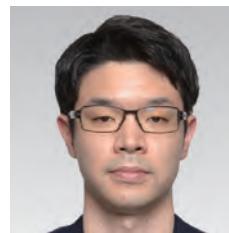
## 参考文献

- 1) 今井：“音声でテレビを楽しむ,” NHK技研R&D, No.175, pp.4-11 (2019)
- 2) World Wide Web Consortium (W3C)：“Web Content Accessibility Guidelines (WCAG) 2.0,”  
<https://www.w3.org/TR/WCAG20/>
- 3) 情報通信研究機構：“情報アクセシビリティの規格,” <https://barrierfree.nict.go.jp/jis/>
- 4) 総務省：“情報バリアフリー環境の整備,”  
[https://www.soumu.go.jp/main\\_sosiki/joho\\_tsusin/b\\_free/b\\_free02.html](https://www.soumu.go.jp/main_sosiki/joho_tsusin/b_free/b_free02.html)
- 5) 一木, 熊野, 今井, 都木：“スポーツ中継番組における自動解説音声の挿入タイミング決定法,” 音響学会春季講演論文集, 2-P-39, pp.1173-1174 (2019)
- 6) M. Ichiki, H. Kaneko, A. Imai and T. Takagi：“A Timing Determination Method for the Insertion of Automated Audio Descriptions into Live TV Sports Programs,” Journal on Technology and Persons with Disabilities, Vol.8, pp.119-129 (2020).
- 7) 一木, 熊野, 今井, 都木：“生放送番組向けの自動解説音声の挿入タイミング決定法について,” 信学技報, Vol.118, No.270, WIT2018-29, pp.45-50 (2018)
- 8) 一木, 宮崎, 今井, 都木：“スポーツ中継番組への解説音声サービスの実証実験における視覚障害者の受容性調査,” 信学技報, Vol.121, No.287, WIT2021-37, pp.28-33 (2021)
- 9) 高橋, 島野, 栗野, 宮崎, 佐野, 今井, 都木：“スポーツ中継を対象とした解説音声制作・配信システム,” 映情学技報, Vol.47, No.7, HI2023-5, pp.109-112 (2023)
- 10) 島野, 栗野, 高橋, 宮崎, 佐野, 今井, 都木：“野球中継における文字認識を用いた解説自動付与,” 第37回人工知能学会全国大会講演論文集, 103-GS-7-01 (2023)
- 11) Y. Kuwano：“AI Application for Enhanced Accessibility of Live Sports Broadcast,” BroadcastAsia 2023 in Asia Tech x Singapore (2023).
- 12) 島野, 栗野, 高橋, 金子：“野球中継映像のみを用いた投球コース推定,” 映情学冬大, 21B-4 (2023)



たかはし まさき  
高橋 正樹

1999年入局。山形放送局を経て、2002年より放送技術研究所にて、画像認識およびアクセシビリティ技術の研究開発に従事。現在、スマートプロダクション研究部チーフ・リード。博士（情報学）。



しまの ゆうき  
島野 雄貴

2016年入局。放送技術局（現メディア技術局）、名古屋放送局を経て、2022年より放送技術研究所にて、画像認識およびアクセシビリティ技術の研究開発に従事。現在、スマートプロダクション研究部所属。



くわの ゆうや  
栗野 裕也

2019年入局。徳島放送局を経て、2021年より放送技術研究所にて、画像認識およびアクセシビリティ技術の研究開発に従事。現在、スマートプロダクション研究部所属。



みやざき まさる  
宮崎 勝

1997年入局。名古屋放送局を経て、2000年より放送技術研究所にて制作支援技術、放送通信連携技術、アクセシビリティ技術、技術と社会との関係構築の研究や企画・広報業務に従事。現在、スマートプロダクション研究部チーフ・リード。技術経営修士（専門職）、博士（工学）。



さの まさひろ  
佐野 雅規

1994年入局。仙台放送局を経て、1997年より放送技術研究所にて、メタデータ制作技術、アクセシビリティ技術の研究開発、国際標準化活動に従事。2023年より、メディア総局メディアイノベーションセンターに所属。チーフ・リード、博士（情報学）。

# 人と一緒にテレビを視聴する ロボットの発話文生成手法

萩尾勇太 上村真利奈 大亦寿之 奥田 誠

## An Utterance Generation Method for TV-Watching Companion Robot

HAGIO Yuta, KAMIMURA Marina, OHMATA Hisayuki and OKUDA Makoto

### 要 約

当所では、人と一緒にテレビを視聴し、番組の内容に基づいて人と対話するコミュニケーションロボット（テレビ視聴ロボット）の研究開発を進めている。本稿では、ロボットが自律的に発話するために、大規模言語モデルを用いて放送番組の字幕文から発話文を生成する手法を提案する。本手法の特徴は、字幕文の文脈を考慮しつつ、字幕文から抽出された各キーワードを話題とした多様な発話文を生成できることである。今回、人が番組の視聴中に話す言葉を想定した発話文を人手で付与した「テレビ番組関連発話データセット」を構築し、大規模言語モデルをファインチューニングすることで発話文生成モデルを開発した。評価実験を行った結果、提案手法は従来手法より自然な文章を生成でき、また番組の文脈に沿った文章を生成できることを確認した。また、特定のキーワードを含んだ、意図した文章を生成できることも確認した。

### ABSTRACT

Our research and development efforts focus on social robots designed for co-viewing television (TV) alongside humans. Here, we propose an utterance generation method for robots that leverages TV subtitles using a large language model. Our method extracts keywords from TV subtitles and generates utterances on the topic of each extracted keyword, considering the context of the subtitles. We constructed a "TV program-related utterance dataset" by manually adding sentences that people might want to say while they are watching TV, and fine-tuned a Japanese large language model. As a result of our evaluation experiments, we confirmed that our method can generate more natural sentences than conventional methods, which accurately align with the ongoing context of the TV program. Additionally, we verified that it could produce sentences that incorporate specific keywords as intended.

## 1. はじめに

家族や友人が同じ空間で一緒にテレビ番組を視聴する“お茶の間視聴”は、番組をきっかけとした会話や気づき・感動の共有など、これまで日常のコミュニケーションを創出する機会の一つとなってきた。一方、近年のスマートフォンや動画配信サービスの普及により、時間や場所に依存しない視聴スタイルが一般的となるなど私たちのメディアへの接触方法は大きく変化している<sup>1)</sup>。また、未婚率の上昇や高齢化などを背景に平均世帯人数は減少しており<sup>2)</sup>、お茶の間視聴が困難な状況となっている。特に、高齢者の社会的なコミュニケーションは認知症リスクを抑制することが知られており<sup>3)</sup>、高齢化が進む日本においてはコミュニケーションの機会を創出することの重要性は、今後ますます高まると考えられる。

このような背景から、当所ではコミュニケーションロボットを活用した新たなお茶の間視聴の実現を目指し、「テレビ視聴ロボット」の研究開発を進めている<sup>4)</sup>。このテレビ視聴ロボットの特徴の一つは、視聴中のテレビ番組の内容に関連したユーザーとの雑談対話ができることである。当所では、この雑談対話に必要なロボットの発話文生成技術の開発を進めてきた。2018年には、視聴中のテレビ番組からキーワードを抽出し、そのキーワードを事前に準備したテンプレート文に埋め込んで発話文を生成する手法を開発した<sup>5)</sup>。開発した発話文生成手法を実装したロボットを用いた評価実験<sup>4)</sup>では、テレビ視聴ロボットの存在により、ユーザーの対話を活性化させる可能性を確認した。一方、従来のロボット発話生成手法<sup>5)</sup>ではロボットの発話文が視聴中の番組の文脈に沿っていないものが多く、似たような番組シーンで同じ発言を繰り返すことがユーザー体験の品質を低下させるという課題が抽出された。

そこで、この課題を解決し、より番組の内容にあった自然な発話を実現することを目的に、大規模言語モデル\*<sup>1</sup>を活用した新しい発話文の生成手法を開発した。評価実験により、本手法を用いることで、番組の文脈を考慮しつつ、番組に含まれる複数のキーワードを含んだ多様な発話文を生成できることを確認した。本稿では、その手法の詳細と評価実験の結果について述べる。

## 2. 既存手法

本章では、テレビ番組の視聴というシーンに限らず、特定の文脈において主体的な発話文を生成する3つの手法を紹介する。

### (1) テンプレート文による文生成法

このアプローチは、事前に準備したテンプレート文にキーワードを埋め込んで文章を生成する手法であり、以前当所にて開発した手法<sup>5)</sup>もその一つである。テレビ番組の映像・音声・字幕文からキーワードを抽出し、事前に準備したテンプレート文の中からそのキーワードに対応するものを選択して組み合わせることで発話文を生成する。この手法により、視聴中のテレビ番組から抽出されたキーワードに関連した発話文を生成できるが、抽出されたキーワードのみを利用していることから、テレビ番組の文脈とかみ合った発話文が生成されるとは限らず、生成される文章も事前に準備したテンプレート文に従った定型的な発話文となることが課題である。

### (2) ソーシャルメディアの利用

このアプローチは、ソーシャルメディアに投稿されたコメントを利用する手法である。Nishimuraら<sup>6)</sup>は、テレビ番組を視聴しているユーザーがソーシャルメディアに投稿したコメントを収集し、それをロボットの発話文としてそのまま利用する手法を提案した。この手法は、実際の視聴者が投稿したテレビ番組の文脈に沿った自然な文章をロボット発話として利用できる利点がある。一方で、テレビ番組に対するコメント数が少ない場合、ロボットの発話数が減少する課題がある。また、不特定のコメントを利用するため、発話の一貫性を保証することが難しいことも課題である。

### (3) 大規模言語モデルの利用

このアプローチは、大規模言語モデルを用いることで、入力されたデータ（文脈）に適した文章を生成する手法である。杉山ら<sup>7)</sup>は、移動中の自動車の車窓画像や周辺スポット情報を基に、ドライブ中の会話を楽しむコミュニケーションロボットの発話文を生成する手法を提案している。ソーシャルメディアから収集した大量の対話データで事前学習した大規模言語モデルを基に、ドライブシーンを対象として収集した対話データでファインチューニング（FT：Fine-tuning）\*<sup>2</sup>することで、ドライブ中の会話に適した発話文を生成するモデルを構築した。想定している利用シーンは異なるものの、テレビ視聴ロボットの発話文生成と同様、時間とともに変化する文脈を考慮した主体的な発話文を生成することを目的と

\*1 大量のテキストデータにより学習した、多量のパラメーターを有するニューラルネットワークモデル。分類や要約、翻訳、テキスト生成など多様なタスクに適用できる高い汎用性を持つ。

\*2 大量のデータを用いて学習したモデル（事前学習済みモデル）に対し、特定のタスクに特化した小規模なデータで最適化する方法。

1表 既存の発話文生成法における利点と課題

	手法	利点	課題
(1)	テンプレート文による文生成法	テレビ番組から抽出されたキーワードごとに複数の発話文を生成できる	テンプレート文に従った定型な発話となり、番組の文脈とかみ合わない文章が生成される
(2)	ソーシャルメディアの利用	番組内容に沿った自然な文章を活用できる	番組に関連した投稿コメントの数や質に依存するため、自動的に文章を生成することはできない
(3)	大規模言語モデルの利用	事前に準備したテンプレート文を用いないため、入力した情報に適した発話文を生成できる	同様の入力情報に対しては似た文章が生成される可能性がある

している。そのため、テレビ視聴ロボットにおいても、視聴しているテレビ番組の情報を入力とすることで、番組の文脈に適した発話文を生成できる可能性がある。

これら3つの手法の利点と課題をまとめたものを1表に示す。本提案手法でこれらのアプローチを組み合わせることで、テレビ番組から抽出された各キーワードに対し、番組シーンに関連したコメントの学習データセットを構築し、大規模言語モデルを用いて各キーワードを話題とした多様な発話文を生成できる手法の確立を目指した。

### 3. 提案手法

本章では、2章で述べたように、テレビ番組から抽出された複数のキーワードそれぞれを話題とし、かつテレビ番組の文脈に沿った発話文を生成する手法を提案する。

#### 3.1 発話文生成

提案する発話生成手法の活用イメージを1図、処理フローを2図に示す。

今回、番組情報としては番組に付与された字幕文を用いることとした<sup>\*3</sup>。発話文生成に字幕文を用いることとした理由は以下の2つである。

- ①事前学習済みモデルの活用: テキストデータの場合、既存の事前学習済みモデルを活用できる。テレビ番組の映像や音声も番組の文脈を理解する上で重要な要素であるが、これらから発話文の生成に寄与する特徴量を得るためには相応の規模を持ったデータセットを準備する必要がある。一方、字幕文はテキストデータであるため、事前学習済みの大規模言語モデルをもとにファインチューニングすることで、比較的小規模なデータセットであっても高精度な発話文の生成を期待できる。
- ②字幕文付与率の向上: 多くの番組で字幕文を用いることが可能である。テレビ番組への字幕文の付与率

は年々向上しており、2022年度の総放送時間に占める字幕放送時間の割合は、NHK総合では90.6%、在京キー5局では69.3%となっている<sup>8)</sup>。

以上の理由から、本稿では字幕データを入力とする発話文生成モデルを提案する。提案手法では、まず発話文生成の対象とする字幕文の中からキーワードを抽出する。キーワードの抽出には、事前に用意したキーワード辞書を用い、本提案手法では固有名詞に強い形態素解析<sup>\*4</sup>のための日本語辞書<sup>9)</sup>に含まれる固有名詞を登録した。2図の例では、字幕文から「アルゼンチン」「ポーランド」「メッシ」の3つのキーワードが抽出される。

次に、抽出された複数のキーワードと字幕文を組み合わせ、発話文生成モデルへの入力データを生成する。入力データは、生成する発話文の話題を指定する [CLS] トークン<sup>\*5</sup>、キーワード、キーワードと字幕文の境界を示す [SEP] トークン、字幕文 (一文ごとに [SEP] トークンを挿入<sup>\*6</sup>) を順に連結して生成する。2図の例では、3つのキーワードが抽出されているため、3つの入力データが生成されている。

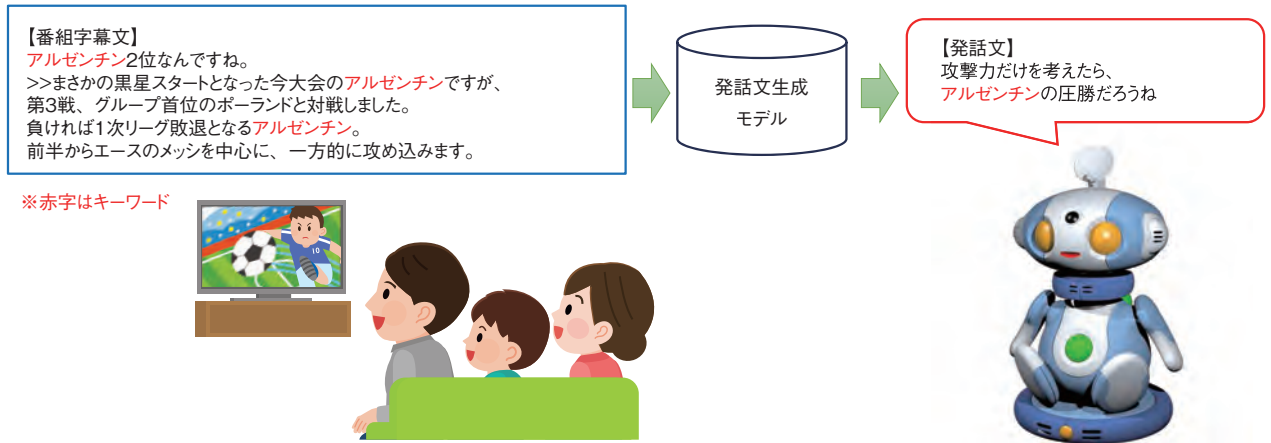
最後に、大規模言語モデルを用いて構築した発話文生成モデルに、生成した入力データを与えることで、入力

\*3 電子番組表 (EPG: Electronic Programming Guide) データに含まれる番組タイトル、概要文、番組ジャンルについても発話文の生成に活用できると考えられる。しかし、EPGは番組全体に関する情報であり、番組の各シーンに関する情報が不足しているため、本提案手法では考慮しないこととした。

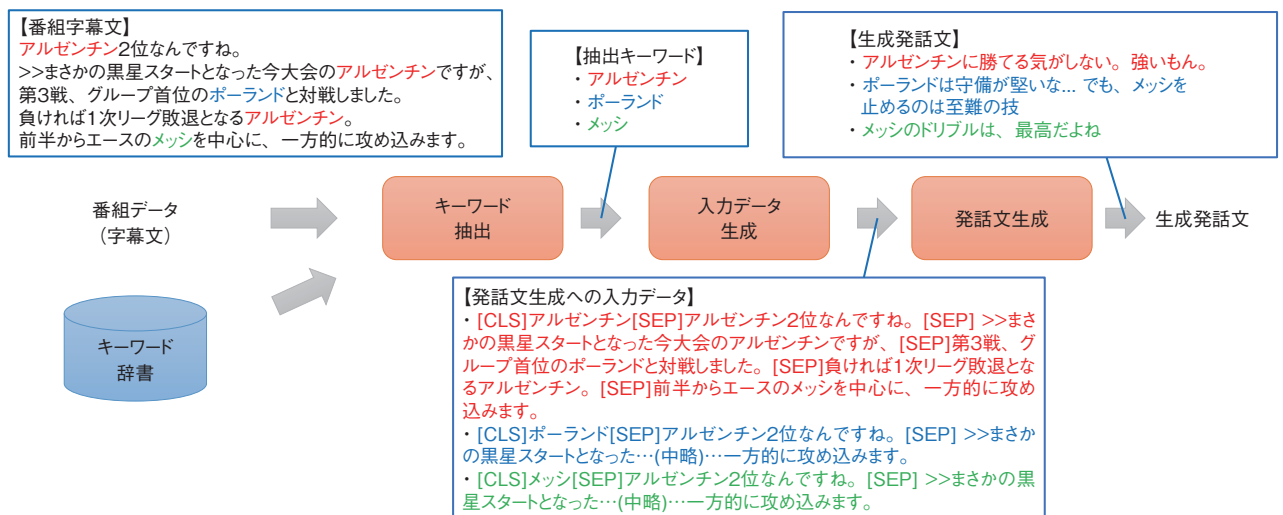
\*4 自然言語処理の手法の一つであり、文法や単語の品詞等の情報を用いて解析し、対象とする文章を形態素と呼ばれる言語において意味を持つ最小単位に分割する処理。

\*5 大規模言語モデルの入出力として用いられるテキストデータの分割単位。文章をトークンへと分割する処理をトークン化処理と呼び、各トークンに一意なIDを割り振ることでモデルへの入出力の文章を数値列として扱う。単語単位や文字単位でのトークン化など様々な手法が存在するが、近年は大量のテキストへの統計処理により効率的な表現を実現するサブワードトークン化がよく用いられる。

\*6 大規模言語モデルが入力テキストの構造を認識しやすくするため、しばしば入力テキストには特殊トークンと呼ばれる記号が用いられる。例えば、文章のクラス分類を示す [CLS] (Classification) トークンや異なる文章の区切りを示す [SEP] (Separator)、トークン列の長さの統一に用いる [PAD] (Padding)、モデルの語彙に存在しない未知の単語を示す [UNK] (Unknown) トークンなどが代表的である。特殊トークンはモデルの学習時と推論時で一貫したパターンで用いることが重要であるが、本提案手法では [CLS] トークンのあとに生成する発話文の話題とするキーワードを指定するものとし、[SEP] トークンによりキーワードおよび字幕文の区切りを表現するものとしている。



1図 提案手法による発話文生成の活用イメージ



2図 提案手法の処理フロー

データに含まれる各キーワードを話題としつつ、番組の文脈に沿ったロボット発話文を生成する。発話文生成モデルは、2図の例で示すように発話文の話題とするキーワードと字幕文を組み合わせた入力データとそれに適した発話文のペアからなるデータセットを構築し、さらにそのデータセットで大規模言語モデルをファインチューニングする。

2図では、「アルゼンチン」を話題とした発話文として「アルゼンチンに勝てる気がしない。強いもん。」、「ポーランド」を話題とした発話文として「ポーランドは守備が堅いな... でも、メッシを止めるのは至難の技」、「メッシ」を話題とした発話文として「メッシのドリブルは、

最高だよ」という発話文が生成されている。これらの発話文は、入力データに埋め込んだキーワードを話題としつつ、字幕文の文脈に沿った発話文となっている。

提案手法により、字幕文から抽出したそれぞれのキーワードを話題とした、多様な発話文を生成することが可能となる。

### 3.2 データセットの作成

3.1節で説明した提案手法を実現するため、キーワードと字幕文を組み合わせた入力データに適した発話文を付与した「テレビ番組関連発話データセット」を作成した。私たちはテレビを視聴しながら番組の内容に基づき

2表 付与された発話文データの例 (NHK総合・首都圏ニュース845)

字幕表示時刻	字幕文	キーワード	ロボット発話文
00:07:39	次のニュースはこちらです。		
00:07:46	JAXAの研究所の近くにある		
00:07:50	相模原市の富士見小学校です。		
		-	小学校どの辺なの？
		JAXA	JAXAの近くなんて羨ましい！
00:07:54	探査機、はやぶさ2が		
00:07:58	小惑星、リュウグウの砂が入って		
00:08:02	いと見られるカプセルを地球へ帰還させる予定の今月6日を前に		
00:08:06	6年生がJAXA フェスを企画しました。		
		-	小惑星にも砂って存在するんだ。
		はやぶさ2	はやぶさ2かっこいい！
		リュウグウ	リュウグウって名前かっこいいな。

3表 構築したデータセットの入出力データの例

入力データ	出力データ
[CLS]NONE[SEP] 次のニュースはこちらです。[SEP]JAXAの研究所の近くにある相模原市の富士見小学校です。	小学校どの辺なの？
[CLS]JAXA[SEP] 次のニュースはこちらです。[SEP]JAXAの研究所の近くにある相模原市の富士見小学校です。	JAXAの近くなんて羨ましい！
[CLS]NONE[SEP] 相模原市の富士見小学校です。[SEP] 探査機、はやぶさ2が[SEP] 小惑星、リュウグウの砂が入って[SEP] いと見られるカプセルを地球へ帰還させる予定の今月6日を前に[SEP]6年生がJAXAフェスを企画しました。	小惑星にも砂って存在するんだ。
[CLS] はやぶさ2[SEP] 相模原市の富士見小学校です。[SEP] 探査機、はやぶさ2が[SEP] 小惑星、リュウグウの砂が入って[SEP] いと見られるカプセルを地球へ帰還させる予定の今月6日を前に[SEP]6年生がJAXAフェスを企画しました。	はやぶさ2、かっこいい！
[CLS] リュウグウ[SEP] 相模原市の富士見小学校です。[SEP] 探査機、はやぶさ2が[SEP] 小惑星、リュウグウの砂が入って[SEP] いと見られるカプセルを地球へ帰還させる予定の今月6日を前に[SEP]6年生がJAXAフェスを企画しました。	リュウグウって名前かっこいいな。

誰かと会話をしたり独り言を言ったりすることがよくあるが、この行動をデータ化したものがこのデータセットである。

大規模言語モデルを用いて字幕文の文脈を考慮した適切な発話を生成させるためには、テレビ番組の実際の字幕文とそれに対応する発話文のペアが必要である。そこで、多様なジャンルのテレビ番組から字幕文を収集し、その字幕文が提示される番組シーンに対し、人が発しそうな文を手で付与することで、データセットを構築した。

まず、2020年から2022年にNHK総合・Eテレで放送された338番組の字幕文を抽出し、字幕文のタイムスタンプと文章を記録したCSV (Comma Separated Values)\*7 ファイルを作成した。作業者はこのCSVファイルの字幕文を確認し、各字幕文から抽出されたキーワードごとにそのキーワードを含んだ発話文を作成した。作業例を2表に示す。

この作業により、計12,396文の発話文を付与した。こ

れらを3.1節で示した入力データと発話文のペアの形式に変換することで、モデル学習用のデータセットを構築した。構築したデータセットの入出力データのペアの例を3表に示す。

### 3.3 モデルの学習

本提案手法では事前学習済みモデルとして、Sugiyamaら<sup>10)</sup>のモデルを利用した。このモデルは、BlenderBot<sup>11)</sup>\*8をベースにソーシャルメディアから取得した約21億ペアの日本語対話対により学習した、約16億パラメーターのTransformer\*9 Encoder-decoderモデル\*10であり、2章で紹介したドライブシーンにおける主体的なロボット発話文生成モデル<sup>7)</sup>の基となったモデルである。

3.2節で構築したデータセットに含まれる計12,396文の入出力データを、学習用と検証用に9:1の比率でランダムに割り当てて、入出力データそれぞれをモデルの事前学習時と同様のトークン化処理<sup>12)</sup>によりトークン列へと変換して学習に用いた。モデルの学習時には杉山



4表 各手法により生成された発話文

字幕文	手法	キーワード	生成発話文
今夜8時からです。今夜には、北海道は広く雪。 あすの朝は西日本で激しい雨の降る所があるでしょう。 このあと、午後には東日本で降り方が強まり、山沿いでは雪になる所もありそうです。 このあと夜には雨雲は海上へと抜けますが、日本海側では雪の降る	提案手法	雪	雪は好きだけど、道路が凍るのは勘弁してほしいよ。
		雨	明日の雨は嫌だなあ。傘を忘れないようにしないと。
	従来手法	雪	雪や関東は今朝と同じか高いでしょう
		雨	まずは強い勢力の雨10号です
	FTなし	-	おはようございます。当地の天気予報は曇りです。今日もよろしくお願いします。
	PTなし	雪	アルゼンチンは、雪がみたいないだいな頑張ったね。
雨		宇宙飛行士は憧れるよね？雪だね。	
どうだ集まったろう。大ごとにするなど言ったではないか。 思ったよりも梶原のおっさん嫌われていたようだ。 (千葉常胤)これは梶原と一戦交えることになりそうだ。戦にはなりませぬ。 加わっていただきかたじけなくございます。 危うくお迎えの支度を始めるところであったわ。	提案手法	梶原	梶原のおっさん、良い人そうだったな。騙したな。
		戦	戦はしちゃいけないのかい？戦はいつでも誰がおこるか分からないよ。
	従来手法	梶原	梶原が本当に大好きでして
		戦	戦本番でも期待したいですね
	FTなし	-	隊長、いきまっせー！大隊長！
	PTなし	梶原	なんだか、お腹が空いてきたな...、面白いな
戦		保険が大切なんだね！意外だな！想像の時期だ！	

ら<sup>7)</sup>の手法を参考に、バッチサイズ<sup>\*11</sup>を64、学習率<sup>\*12</sup>を $1.0 \times 10^{-5}$ で3,000ステップ<sup>\*13</sup>学習し、検証データによる推論時の損失が最も小さいモデルを利用した。

## 4. 評価実験

提案手法を評価するため、生成した発話文の自然さと正確性についてアンケート調査を用いた実験を行った。

### 4.1 実験の概要

2022年にNHK総合で放送された5つの番組に含まれる字幕文の中から、連続した5つの文から構成される番組シーン3,152個を作成した。それらの中から各番組2個ずつ、計10個の評価用番組シーンを無作為に選択して評価実験に用いた。なお、評価用番組シーンの生成に用いた番組は、提案手法の発話文生成モデルの訓練データおよび検証データに含まないものであり、「ニュース／報道」「情報／ワイドショー」「ドラマ」「バラエティー」の4つのジャンルの番組を含んでいる。

各評価用番組シーンに対して、以下の4つの手法により発話文を生成し、アンケート回答者に提示した。

- 提案手法：3章にて説明した方法で構築した発話文生成モデルを用いる。入力された字幕文からキーワードを抽出し、各キーワードを話題とした発話文を生成する。
- 従来手法<sup>5)</sup>：入力された字幕文からキーワードを抽出し、事前に準備したテンプレート文の中から、抽出さ

れた各キーワードに適した文を選択し、組み合わせることで発話文を生成する。

- FT (Fine-tuning) なし<sup>10)</sup>：3.2節で説明したテレビ番組関連発話データセットでファインチューニングする前の大規模言語モデルにより発話文を生成する。
- PT (Pre-training) なし：事前学習済みモデルを用いず、テレビ番組関連発話データセットのみで学習したモデルにより発話文を生成する。

FTなしはテレビ番組関連発話データセットによるファインチューニングの効果を検証するため、PTなしは大量の日本語文による事前学習の効果を検証するために比較対象とした。なお、評価用番組シーンから提案手法により抽出されたキーワード20個のうち、訓練データ

- \*7 値や項目をカンマ(,)で区切るにより表形式で表現するデータ構造。表計算ソフトで扱いやすい特徴がある。
- \*8 Meta Platforms社が開発し、2020年に公開したオープンソースのチャットボットモデル。人格や知識の一貫性、相手への共感といった会話に必要な複数のスキルを学習し、融合させることで自然な対話を実現している。
- \*9 自己注意機構と呼ばれる構造を導入することで、自然言語処理で従来用いられてきた再帰的ニューラルネットワークと比較して、長い入力文に対する処理能力と計算効率を向上させたモデル。
- \*10 入力文を特徴量に変換するエンコーダーと特徴量から出力文に変換するデコーダーの構造を組み合わせたモデルであり、主に異なる言語間の翻訳タスクに用いられる。
- \*11 機械学習モデルの学習時に一度に処理するデータの数。モデルのパラメーター数や利用できる計算機資源に応じて設定する必要がある。
- \*12 モデルのパラメーターを更新する際に、一度に変動させる量を調整する値。学習速度とモデルの精度を両立させる値に設定することが重要であり、学習のフェーズに応じて調整することも行われる。
- \*13 モデルのパラメーターを更新する回数。回数が多いほどモデルの精度は向上するが、多すぎると学習データに含まれていないデータへの推論性能が低下する過学習と呼ばれる状態になるため、モデルおよび学習データの規模に応じた適切な値に設定する必要がある。

に含まれていた単語は4個のみであり、評価に用いたキーワードの多くは訓練データに含まれていないものであった。

前述の4つの手法により生成された各発話文をアンケート回答者に提示し、以下の3つの設問に回答してもらった。その際、アンケート回答者が提示された発話文の生成手法を意識しないように、発話文は設問ごとにランダムな順序で提示した。

Q1. 発話文は、日本語として自然な文章となっているか

Q2. 発話文は、提示した字幕文の番組シーンに合った内容か

Q3. 発話文は、キーワードを話題としたものとなっているか

Q1とQ2は4つの手法すべてを対象とした5段階のリッカート尺度<sup>\*14</sup>で回答してもらい、Q3は提案手法と従来手法を対象とした3段階のリッカート尺度で回答してもらった。アンケートの回答者は4名であり、各人が合計70文を評価した。各手法により生成された発話文の例を4表に示す。

## 4.2 実験結果

アンケート調査による評価結果を3図～5図に示す。

3図は、発話の日本語としての自然さに関する質問項目(Q1)の結果である。提案手法は約88%が「全く違和感がない」「ほぼ違和感がない」との回答であったのに対し、従来手法では約62%、FTなしでは約44%、PTなしでは約20%であった。したがって、提案手法は従来手法と比較して、日本語としての発話文の自然さが約26ポイント向上した。

4図は、入力された字幕文の文脈上の自然さに関する質問項目(Q2)の結果である。提案手法は約75%が「全く違和感がない」「ほぼ違和感がない」との回答であったのに対し、従来手法では約24%、FTなしでは約20%、PTなしでは約4%であった。したがって、提案手法は従来手法と比較して、字幕文の文脈に対する発話文の自然さが約51ポイント向上した。

5図は、生成された発話文がキーワードを含む妥当な文章となっているかに関する質問項目(Q3)の結果である。提案手法では約99%が「話題にしている」であったのに対し、従来手法では約90%であった。したがって、提案手法は従来手法と比較して、字幕文から抽出されたキーワードに対する発話文の関係性を約9ポイント向上した。

## 4.3 実験結果の考察

4.2節に示した評価実験の結果から、提案手法が従来手法と比較して、日本語としても、字幕文の文脈としても、より自然な発話を生成できることを確認した。これは、3.2節にて構築したテレビ番組関連発話データセットを用いて、大規模言語モデルをファインチューニングしたことにより、字幕文の文脈を考慮した発話文の生成能力が向上したためと考えられる。

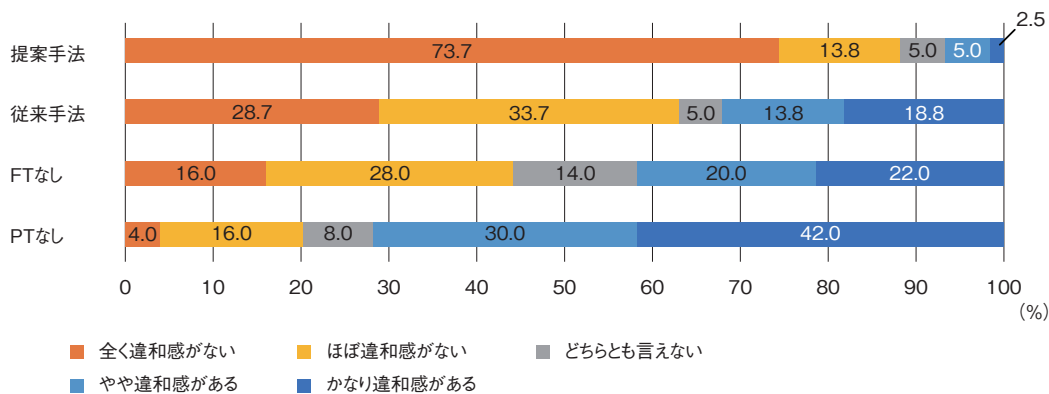
また、FTなしおよびPTなしにより生成された発話文は、Q1とQ2ともに従来手法よりも評価が低いことを確認した。このことから、視聴中のテレビ番組に適した発話文を生成するというシーンにおいても、大量の日本語テキストによる事前学習および特定のタスクに特化したデータセットによるファインチューニングが有効であることを示している。

一方、提案手法においても、ドラマなどの字幕文のみでは番組の文脈を推定しづらい場面では評価が低い傾向があった。今回は字幕文のみを入力として発話文の生成を試みたが、電子番組表に含まれる番組タイトルや概要文も含めてモデルを学習することで、より精度の高い発話文を生成できる可能性がある。

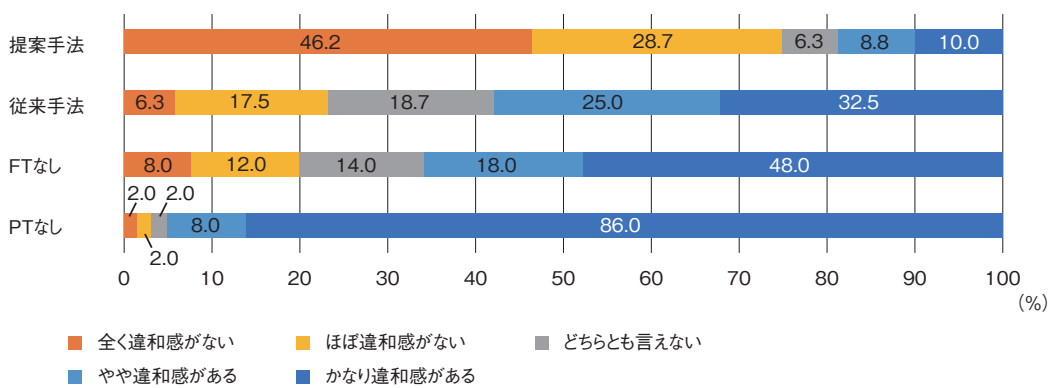
また、提案手法は、入力された一つの番組シーンに対して、抽出されたキーワードそれぞれを話題とした多様な発話文を生成できることを確認した。しかし、生成した複数の発話文から実際にロボットが発話する文章を選択する手法については課題として残されている。当所で開発した、個人のテレビ番組の視聴履歴からそのユーザーの興味を推定する手法<sup>13)</sup>を用いることで、ユーザーの興味度が高いキーワードを話題とした発話文を選択できると考えている。

以上より、提案手法はテレビ番組の内容に基づいた雑談対話を行うテレビ視聴ロボットの発話文を生成する手法として一定の性能があることを確認した。

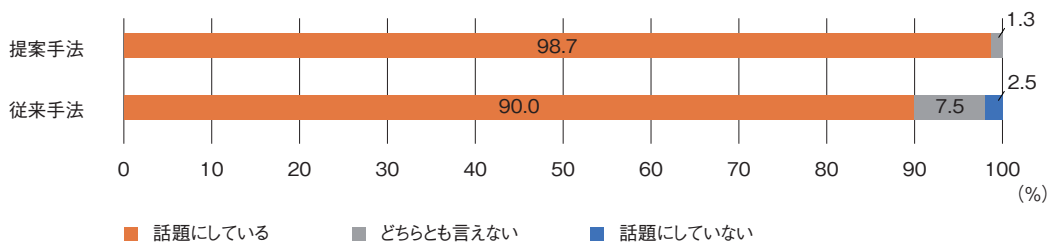
\*14 アンケート調査などで用いられる回答尺度の一つであり、提示された質問に対してどの程度合意できるかを回答として求める。一般的に5段階や7段階の尺度を用いることが多い。



3図 「Q1. 発話文は、日本語として自然な文章となっているか」の評価結果



4図 「Q2. 発話文は、提示した字幕文の番組シーンに合った内容か」の評価結果



5図 「Q3. 発話文は、キーワードを話題としたものとなっているか」の評価結果

## 5. おわりに

本稿では、変化するメディア環境と社会構造の中で機会が減少しつつあるお茶の間視聴に代わる新しいテレビ視聴体験の実現に向けて、当所で研究開発を進めているテレビ視聴ロボットの発話文の生成手法を提案した。番組の文脈を考慮した自然な発話を実現するため、字幕文から発話文を生成する過程において大規模言語モデルを用いることとした。具体的には、テレビ番組の視聴中を想定した発話文により構成される「テレビ番組関連発話データセット」を構築し、大規模言語モデルのファイン

チューニングを行った。

評価実験の結果、提案手法は従来手法と比較して、日本語の自然さ、番組の文脈上の自然さ、キーワードを正しく話題として取り上げる点において、性能が向上することを確認した。これらの結果から、テレビ視聴ロボットの発話文生成において、提案手法の有効性を示した。

一方、ドラマなどの複雑な文脈を持つ番組において、字幕文のみから番組の全体的な文脈を把握することは難しく、番組に適した発話文が生成されないことも課題として確認した。今後は、番組の概要やジャンル情報など、字幕文以外の情報源も活用した手法を検討したい。

## 参考文献

- 1) 総務省：“令和4年版情報通信白書,” (2022)
- 2) 厚生労働省：“国民生活基礎調査の概況,” (2021)
- 3) T. Sato, C. Murata, M. Saito, T. Takeda and K. Kondo: “Influence of Social Relationship Domains and Their Combinations on Incident Dementia: a Prospective Cohort Study,” J. Epidemiology Community Health, Vol. 72, No.1, pp.7-12 (2018).
- 4) Y. Hagio, M. Kamimura, Y. Hoshi, Y. Kaneko and M. Yamamoto: “TV-Watching Robot: Toward Enriching Media Experience and Activating Human Communication,” SMPTE Motion Imaging J., Vol.131, No.4, pp.50-58 (2022).
- 5) 金子, 星, 村崎, 上原：“テレビ番組に含まれるキーワードに基づく発話文生成手法,” 情処研報, Vol.2018-IFAT-131, No.2, pp.1-8 (2018)
- 6) S. Nishimura, M. Kanbara and N. Hagita: “Atmosphere Sharing with TV Chat Agents for Increase of User’s Motivation for Conversation,” International Conference on Human-Computer Interaction, Human Aspects of IT for the Aged Population, pp.482-492 (2019).
- 7) 杉山, 古賀, 西島：“移動体から見える風景を話題とする雑談対話システム,” 人工知能学会全国大会論文集, 2N5-OS-7a-04, pp.1-4 (2022)
- 8) 総務省：“令和4年度の字幕放送等の実績” (2023)
- 9) 佐藤, 橋本, 奥村：“単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討,” 言語処理学会 第23回年次大会, pp.875-878 (2017)
- 10) H. Sugiyama, M. Mizukami, T. Arimoto, H. Narimatsu, Y. Chiba, H. Nakajima and T. Meguro: “Empirical Analysis of Training Strategies of Transformer-based Japanese Chat Systems,” IEEE Spoken Language Technology Workshop, pp.685-691 (2023).
- 11) S. Roller, E. Dinan, N. Goyal, D. Ju, M. Williamson, Y. Liu, J. Xu, M. Ott, K. Shuster, E. M. Smith, Y. Boureau and J. Weston: “Recipes for Building an Open-Domain Chatbot,” Proc. 16th Conf. the European Chapter of the Association for Computational Linguistics: Main Volume, pp.300-325 (2021).
- 12) T. Kudo and J. Richardson: “SentencePiece: A Simple and Language Independent Subword Tokenizer and Detokenizer for Neural Text Processing,” Proc. 2018 Conf. Empirical Methods in Natural Language Processing: System Demonstrations, pp.66-71 (2018).
- 13) 奥田, 金子, 星, 萩尾, 上村, 西本：“電子番組表解析によるテレビ視聴者の興味推定,” 映情学技報, Vol.46, No.13, pp.33-35 (2022).



はぎお ゆうた  
萩尾 勇太

2015年入局。福島放送局を経て、2018年から放送技術研究所において、人と一緒にテレビを視聴するロボットの研究に従事。現在、放送技術研究所ネットサービス基盤研究部に所属。



かみむら まりな  
上村 真利奈

2016年入局。仙台放送局を経て、2019年から放送技術研究所において、人と一緒にテレビを視聴するロボットの研究に従事。現在、放送技術研究所ネットサービス基盤研究部に所属。



おくた ひさゆき  
奥田 誠

2003年入局。甲府放送局を経て、2007年から放送技術研究所において、放送通信連携システムの研究開発と標準化に従事。技術局を経て、現在、放送技術研究所ネットサービス基盤研究部チーフ・リード。



おくた まこと  
奥田 誠

2000年入局。大阪放送局を経て、2003年より放送技術研究所にて、機械学習、画像処理、言語処理、ロボット工学に関する研究に従事。2015年から2021年まで（国研）情報通信研究機構に。現在、放送技術研究所ネットサービス基盤研究部副部長。

## コンテンツ発見技術 ～“放送”への多様なアクセスの実現を目指して～

NHKの番組を視聴したいとき、放送中であればテレビのチャンネルを合わせて、放送時間が過ぎてしまった場合はテレビやスマートフォンのNHKプラスのアプリで見ることができる。このように、いつでも・どこでも見たい番組を見られるようになったが、その一方で同じ番組にもかかわらず、視聴する環境や放送時間に合わせてデバイスやアプリを選んで操作しなければならない。また、新たなデバイスやアプリが登場すれば番組を視聴する機会は増えるが、操作方法がそれぞれ異なると、緊急時などに慣れない操作に手間取り、重要な情報を見逃してしまうかもしれない。

当所では、このような視聴方法の多様化に伴う煩雑さや不便さを無くすために、場所や時間にとらわれずに簡単に統一された視聴体験を提供する“コンテンツ発見技術”の研究に取り組んでいる。

### ■ デバイスの機能や番組の提供状況の違いを意識させない視聴体験

コンテンツ発見技術を用いて、例えば共通のWebアプリ（1図）がデバイスの放送受信機能の有無や放送・同時配信・オンデマンド配信などの番組の提供状況に応じた視聴方法を自動で判断することで、デバイスによらず同じ操作で番組を視聴できるようになる。

ここで重要となるのが、Webアプリが参照するコンテンツ発見サーバーのメタデータである。番組の詳細情報（タイトル、ジャンルなど）と提供情報（放送チャンネル、配信URLなど）を分離・構造化することで、適切な方法で番組にアクセスすることができる（2図）。

### ■ Web標準\*1に合わせた番組メタデータの共通化

テレビの電子番組表（EPG）のメタデータは放送の標準規格に準じているが、ネット上で使用されるメタデータの多くはWeb標準である“Schema.org\*2”形式で記述されている。コンテンツ発見技術で扱う番組のメタデータについては、ネット上の情報との関連づけを行いやすくするため、同様の形式を採用した。これにより、例えば教育や小売など多種多様なサービスと放送とのスムーズな連携が期待される。

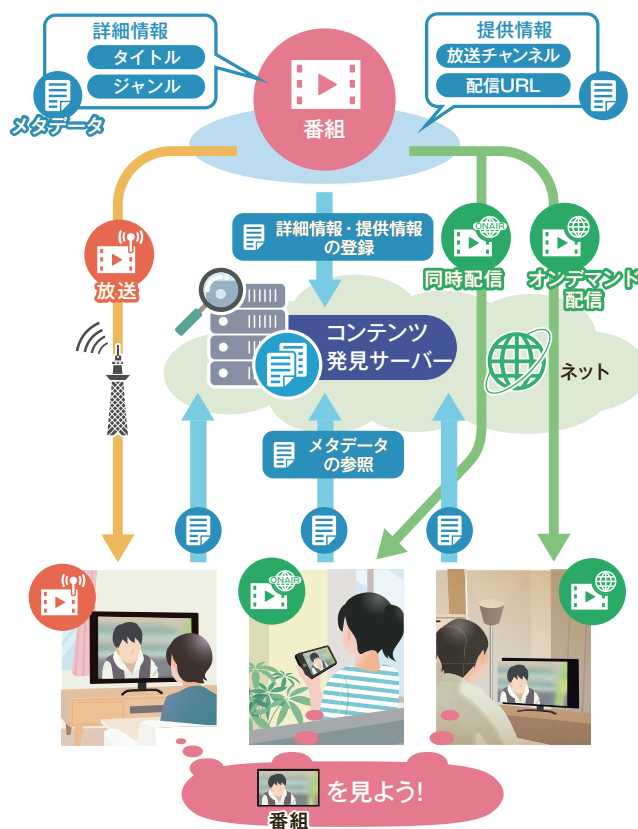
今後は、さまざまな業界と協調しながらコンテンツ発見技術の検証や標準化を進めていく。

\*1 Webの国際標準化団体であるW3C（World Wide Web Consortium）で策定された仕様やガイドライン。

\*2 ネット上で標準的に利用されている構造化データの仕様。



1図 コンテンツ発見技術を用いたWebアプリ（試作）



2図 コンテンツ発見技術による新しい視聴体験のイメージ

# 長寿命で省電力な有機ELディスプレイを目指して

当所では、ディスプレイの利便性を高めるため、薄くて軽く柔らかいフレキシブルな有機EL<sup>\*1</sup>の研究開発を進めている。これが実現すると、大画面のディスプレイを小さく収納したり、さまざまな形状に仕立てたりできるため、新しいスタイルでの視聴が可能になる。

## ■フレキシブル有機ELディスプレイの長寿命化

フレキシブル有機ELディスプレイを実用化するための課題の1つが有機ELの寿命である。ディスプレイを薄くて軽く柔らかくするには、薄いフィルムを基板に用いる必要がある。しかし、薄いフィルムは大気中の酸素や水分を通しやすいため、これらに弱い材料を用いた有機ELでは、長時間の発光により材料が劣化して全体的に明るさが低下したり、光らない部分が生じて発光面が虫食い状になったりする。スマートフォンなどで使われる小型の有機ELディスプレイでは、酸素や水分を遮断するために高価な「封止膜」を使用している。当所では、フレキシブルディスプレイの普及に向けて、酸素や水分を透過しやすい安価なフィルム上でも長持ちする有機EL素子を開発した。

有機EL素子は、陽極から正孔が、陰極から電子が注入され、発光層で電荷が再結合することで発光する(1図)。当所では、陰極から電子を供給するための電子注入層として、従来のリチウム系の材料と比べて酸素や水分に強い有機材料を開発するとともに、陽極から正孔を供給する正孔注入層として酸素や水分に強い材料構成を見だし、フィルム上でも長持ちする有機EL素子を実現した(2図)。また、これらの材料を照明に適用することで、厚さ0.07mmの非常に薄く柔軟な照明を開発した(3図)。これにより、照明装置の自由度向上が期待できる。

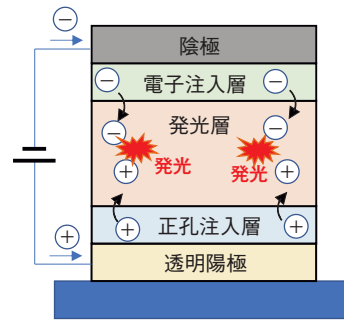
## ■省電力な有機EL素子の実現

開発した材料を用いると、酸素や水分に対する耐久性だけでなく、電子や正孔の供給しやすさも同時に改善できる。従来材料を用いた素子では十分な光を得るのに数ボルトの電圧が必要であったが、今回、乾電池1本(1.5ボルト)でも十分な明るさで光る省電力な青色有機EL素子を実現できた(4図)。

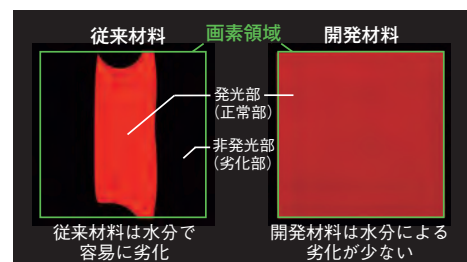
今後は、さらなる長寿命化、省電力化を目指すとともに、照明や太陽電池などの有機デバイスへの応用展開も進めていく。

本研究は、(株)日本触媒と共同で進めている。

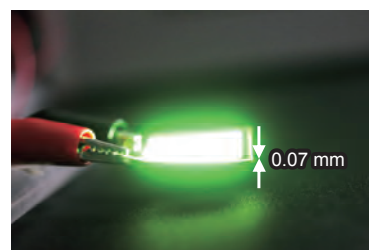
\*1 有機エレクトロルミネッセンス (Electro-Luminescence) の略。有機化合物を用いた層状の構造体に電流を流すと発光する現象。



1図 有機EL素子の構造



2図 水分透過下で長時間発光させた有機EL素子の比較



3図 開発した照明



4図 開発した省電力な有機EL素子

## 3次元空間の撮影位置自動選択技術

近年の映像制作では、3Dコンピューターグラフィックス（3DCG）の活用が、より一般的になっている。3DCGを活用した映像制作では、複数の面で構成される3Dオブジェクト（以下、オブジェクト）を仮想空間に配置し、コンピューターの中に構成した3Dシーン（以下、シーン）を仮想カメラで撮影する。

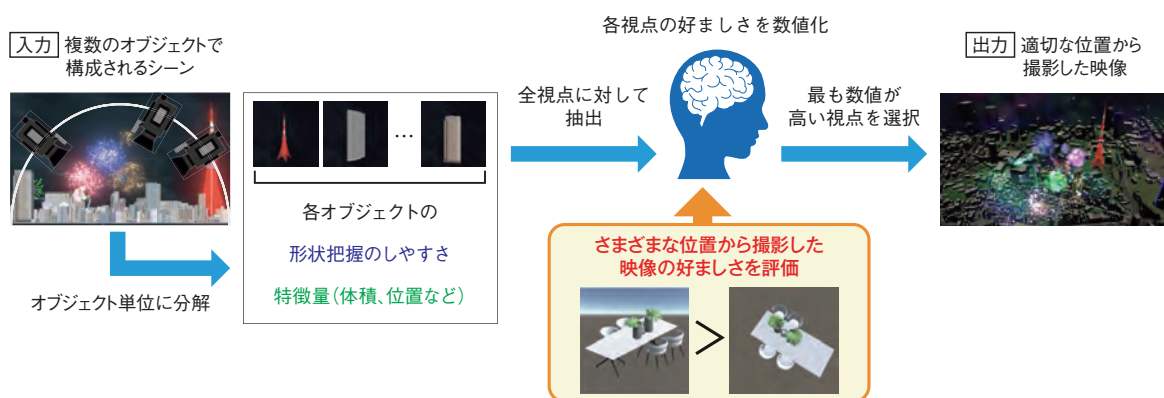
仮想カメラは、物理的な制約を受けずに自由な位置に配置できるため、実世界を撮影する場合と比べて、撮影位置の候補が膨大となる。これらの中から、制作者の意図に合った仮想カメラの位置を手動で選択する作業は、手間と時間がかかる。これを自動で選択できれば映像制作の効率化ができるため、当所では、適切な撮影位置を自動で選択するアルゴリズムの開発に取り組んでいる。

### ■ 適切な撮影位置を自動で選択するアルゴリズム

物体やシーンを観察する際に、人が好ましいと考えて選択する視点位置は、斜め上など形状の複雑さを把握しやすい位置であることが知られている。これまで、オブジェクトを構成する複数の面のうち、視点から見える面の数や面積を基に形状把握のしやすさを数値化し、人が好む視点を推定する“視点選択アルゴリズム”が複数提案されている。しかし、これらは単一のオブジェクトで構成されるシーンを想定しており、映像制作で扱うような多数のオブジェクトから構成されるシーンは想定していなかった。

当所では、多数のオブジェクトから構成されるシーンにも適用可能な、新しい“撮影位置選択アルゴリズム”を開発した（1図）。このアルゴリズムは、さまざまな位置から撮影した映像の好ましさを人が評価したデータセットを使用して、オブジェクトの体積や位置などの特徴量が映像の好ましさに影響を与える度合い（特徴重要度）を学習している。まず、シーンをオブジェクト単位に分解し、既存の視点選択アルゴリズムを用いて、全視点に対して形状把握のしやすさおよび特徴量を算出する。次に、算出した形状把握のしやすさおよび特徴量に、特徴重要度を掛け合わせることで、各視点の好ましさを数値化する。最後に、全視点の中で最も数値が高い視点を選択し、その位置から撮影した映像を出力する。

今後は、推定できる撮影位置の範囲を広げてアルゴリズムの精度向上を図り、より効率的な映像制作の実現を目指す。



1図 開発した撮影位置選択アルゴリズムの処理の流れ



## 多元系半導体量子ドットを使った高色純度EL素子の開発

映像情報メディア学会誌, Vol.78, No.1, pp.124-130 (2024)

本村玄一, 大久 哲, 上松太郎<sup>\*1</sup>, 桑畑 進<sup>\*1</sup>, 亀山達矢<sup>\*2</sup>, 鳥本 司<sup>\*2</sup>, 都築俊満, 藤崎好英

※1 大阪大学 ※2 名古屋大学

高画質かつ多様な形態のディスプレイを実現できる技術として、量子ドットを用いた発光素子の開発に取り組んでいる。量子ドットは粒子サイズや材料組成で光学特性を制御できる結晶微粒子で、色純度の高い発光が得られることから、色再現性に優れた広色域ディスプレイに利用されている。硫化銀インジウムなどに代表される多元系半導体（複数の元素から構成される化合物半導体）からなる量子ドット材料の発光波長制御に取り組み、材料組成に応じて高色純度かつ多様な発光色を示す量子ドット材料を得た。合成した各色の多元系半導体量子ドットを用いて、電流を流して発光（EL：Electroluminescence）させる量子ドットEL素子を試作した。色純度の低下につながるEL素子に特有の欠陥発光を抑制し、各色共通の素子構造で色純度の高いELを実現したことで、多元系半導体量子ドットによるEL素子が広色域ディスプレイに有望であることが示された。

## Design for Stray Light Suppression of an Optical Phased Array with Electro-optic Polymer

ITE Transactions on Media Technology and Applications, Vol.12, No.1, pp.93-99 (2024).

宮本裕司, 三浦雅人, 町田賢司, 上田里永子<sup>\*</sup>, 山田千由美<sup>\*</sup>, 山田俊樹<sup>\*</sup>, 大友 明<sup>\*</sup>, 平野芳邦

※ 情報通信研究機構

3次元映像表示用ディスプレイや距離センサーへの応用を目指して、光ビームの向きを高速に変えることができる、電気光学（EO：Electro-Optic）ポリマーを用いた導波路型光フェーズドアレイ（OPA：Optical Phased Array）の研究を進めている。導波路における光の伝搬では、特に曲げ導波路や光スプリッターにおいて損失が生じやすく、損失により漏れた光は迷光となる。迷光はディスプレイの画質劣化や距離センサーのSNR（Signal-to-Noise Ratio）低下の原因になるため、曲げ導波路における放射損失や、光スプリッターにおける結合損失を抑制した導波路設計が重要である。本研究では、曲げ導波路の曲線部、および光スプリッターの入力部を通過する光の強度分布変化に着目した。また、導波路における光入出力パワー比についてビーム伝搬法を用いた数値シミュレーションによって算出することで、光の損失を抑制できる最適設計手法を提案した。さらに、本手法により設計したOPAについて動作シミュレーションおよび試作素子を用いた評価実験を行い、迷光の抑制および光ビームの鮮鋭度の改善効果を確認した。

## Demonstration Experiment of a System that Helps Users Select Educational Content Using Terms in Basic/advanced Relationships

Journal of Information Processing, Vol.31, pp.766-774 (2023).

藤井翔子, 阿部晋矢, 藤津 智, 松村欣司, 藤沢 寛

インターネット上ではさまざまな学習プラットフォームが利用できるようになり、動画コンテンツの教育利用が進む中、オンтоロジーを利用した教育・学習支援に関する研究も盛んに行われている。しかし、これらの多くの場合、コンテンツの体系化は教科内に限定されており、教科横断的な連携や、教科を超えた学習者の興味関心の広がりに関する検討はされていない。本論文では、教科の枠を超えてあらゆる教育コンテンツを体系化することを目指し、人が学習を進めていく過程での単語間の学習順序の関係性を基礎／発展関係として新たに定義し、この関係下にある単語対を構造化・活用することで、教育コンテンツ間の学習順序を推定する手法を提案した。また、提案手法を組み込んだ教育コンテンツ選択支援システムを試作し、検証実験を行った。教科書の順序に基づいたコンテンツ提示手法と、提案する教科横断的なコンテンツ提示手法とを比較することで、各手法の特徴を明らかにし、提案手法における学習者の興味関心の広がりへの有効性を確認した。

## Capturing Videos at 60 Frames per Second Using Incoherent Digital Holography

Optics Continuum, Vol.2, No.11, pp.2409-2420 (2023).

室井哲彦, 信川輝吉, 片野祐太郎, 萩原 啓

高精細な3次元動画の撮影を目指して、インコヒーレントデジタルホログラフィー（IDH：Incoherent Digital Holography）の研究を進めている。IDHは太陽光やLED照明光などのインコヒーレント光のもとでホログラムを生成できるため、従来のカメラと同じ使用環境で3次元動画を取得できる。IDHで動画を取得する方法として、2つの焦点距離を持つ二重焦点レンズと偏光カメラを用いた光学系による空間分割位相シフト法が提案されている。二重焦点レンズを回折光学素子で構成する場合、高次回折光や散乱光が生じるため、取得するホログラムにノイズが重畳する。そこで、二重焦点レンズと同じ機能を有するマイケルソン干渉計を構築し、IDH光学系に導入した。最大60fpsで動画を取得できることを実証し、動画像ばやけが従来のカメラと同程度であることを確認した。これにより、IDHは放送と同じフレームレートで3次元動画を取得できることを明らかにした。

公開番号	発明考案の名称	技術概要
特開2023-170185	類似領域検出装置及び類似領域検出プログラム	探索領域の位置や形状によらず一定の計算量で、画像の特徴量のヒストグラムが類似する領域を検出することで、目標画像に類似する、入力画像内の類似領域を検出することを可能とする類似領域検出装置およびそのプログラム。
特開2023-170822	解説音声制作装置及びプログラム	複数の情報源のデータを利用するとともに、拡張性および汎用性の高い解説音声をリアルタイムで提供可能な解説音声制作装置およびプログラム。
特開2023-171329	受信装置およびプログラム	放送サービスあるいはインターネットなどでのネット配信サービスに関するサービス情報を適切に取得すること、放送サービスとネット配信サービスを統合した抽象サービスを生成すること、しかるべきタイミングにおいて上記のサービス情報の更新を適切に行うこと、の少なくともいずれかを行うことのできる受信装置およびプログラム。
特開2023-171341	受信装置およびプログラム	放送サービスあるいはインターネットなどでのネット配信サービスに関するサービス情報を適切に取得すること、放送サービスとネット配信サービスを統合した抽象サービスを生成すること、しかるべきタイミングにおいて上記のサービス情報の更新を適切に行うこと、の少なくともいずれかを行うことのできる受信装置およびプログラム。
特開2023-172942	受信装置およびプログラム	放送サービスあるいはインターネットなどでのネット配信サービスに関するサービス情報を適切に取得すること、放送サービスとネット配信サービスを統合した抽象サービスを生成すること、しかるべきタイミングにおいて上記のサービス情報の更新を適切に行うこと、の少なくともいずれかを行うことのできる受信装置およびプログラム。
特開2023-173001	手話単語時間長算出装置およびそのプログラム、ならびに、手話CG映像生成装置およびそのプログラム	一連の手話動作における個々の手話単語の時間長を算出することが可能な手話単語時間長算出装置およびそのプログラム、ならびに、手話単語ごとの時間長に対応して手話CG映像を生成することが可能な手話CG映像生成装置およびそのプログラム。
特開2023-173088	スピナー制御装置およびそのプログラム	1度に増減させる増減量を調整可能として、スピナーの増減操作を行うことができ、また、増減方向提示領域によって、スピナーの増減方向を視覚化することができるスピナー制御装置およびそのプログラム。
特開2023-173602	有機エレクトロルミネッセンス素子、表示装置および照明装置	高い発光効率を有し、寿命の長い有機エレクトロルミネッセンス素子、そのような有機エレクトロルミネッセンス素子を備える表示装置および照明装置。
特開2023-173608	有機エレクトロルミネッセンス素子、表示装置および照明装置	高い発光効率を有し、寿命の長い有機エレクトロルミネッセンス素子、そのような有機エレクトロルミネッセンス素子を備える表示装置および照明装置。
特開2023-173902	撮像素子	共有画素構造1つあたりに追加するトランジスタを1個に抑制して小型化を図りつつ、各制御単位ごとに、空間解像度と時間解像度の2つの要素をより細かく制御し得る構造を備えた撮像素子。
特開2023-173975	顔認識装置及びそのプログラム	顔認識において、認識対象者の識別情報と登録者の各顔特徴量とを対応付けたハッシュ型マップテーブルを用いて、同一の登録者についての重複部分を除去することで、計算量を抑制できる顔認識装置およびそのプログラム。
特開2023-174033	立体配線を用いた回路及びその作製方法	フィルム基板の表面側配線と立体配線との間で、電気的な接続不良の発生やコンタクト抵抗の増加を抑制できる回路およびその作製方法。
特開2023-176192	撮像素子の製造方法	入射した光が撮像素子の透明基板内で隣接する画素に漏れこみ、混色(クロストーク)が生じることを抑制できる撮像素子の製造方法。
特開2023-177648	興味推定装置、推薦装置及びプログラム	ユーザーが興味を持っている単語を、精度よく検出することが可能な興味推定装置、推薦装置およびプログラム。
特開2023-177676	音響メタデータ座標系変換装置及びプログラム	任意のスピーカー配置に対して、各スピーカーの極座標値から規格値と一致する直交座標値へ変換、または直交座標値から規格値と一致する極座標値へ変換でき、かつ音声オブジェクトの座標値のずれを補正することが可能な音響メタデータ座標系変換装置およびプログラム。

公開番号	発明考案の名称	技術概要
特開2023-178045	感覚提示制御装置及びそのプログラム	VR映像と連動して温感を提示する感覚提示制御装置およびそのプログラム。
特開2023-179125	デジタル音声信号同期装置及びプログラム	異なる同期系で動作している複数のシステムの組み合わせにより構成されている大規模な設備において、一のシステムが非PCM (Pulse Code Modulation: パルス符号変調) のデータを含む異なる同期系のデジタル音声信号を受信した場合であっても、ビット単位で前記デジタル音声信号のデータが保持されている状態で、受信したシステムの同期に引き込むことができるデジタル音声信号同期装置およびプログラム。
特開2023-179298	半導体デバイス及びその製造方法	トランジスターの安定動作が可能な、柔軟性を備えた半導体デバイスおよびその製造方法。
特開2023-179648	符号化装置、復号装置、及びプログラム	映像符号化においてACT(Adaptive Colour Transform: 適応色変換) を適用する場合であっても、画質の劣化を抑制可能な、符号化対象ブロックごとに符号化する符号化装置、復号装置およびプログラム。
特開2023-179855	映像分割装置、映像合成装置及びプログラム	任意のアスペクト比および画素数の映像信号を伝送することが可能な映像分割装置、映像合成装置およびプログラム。
特開2023-179864	視点変換装置及びそのプログラム	立体映像の視点を変換する際、余白による構図を維持できる視点変換装置およびそのプログラム。
特開2023-180280	光バイパススイッチ	光バイパススイッチを用いたマルチキャスト伝送系における特定経路の保護 (特定経路中のルーターの停電や故障などの障害からの特定経路の保護) について、障害からの回復時にルーターを復帰させる際のマルチキャストストリームの遮断を回避することができる光バイパススイッチ。
特開2023-180550	放送システム、中継装置、及び算出装置	設備コストを増大させたり送信信号の改変を伴ったりせずに、中継装置間における放送信号の送信時間差をより高い精度で測定することを可能にする放送システム、中継装置、および算出装置。
特開2023-181735	磁性細線メモリ	同じ磁化方向の連続した磁区であっても、安定した一定長の磁区を磁性細線に形成することができる磁性細線メモリ。
特開2023-181813	伸縮性を有する配線基板及びその製造方法	高い伸縮性と導電性とを兼ね備えた導電性薄膜によって、微細な配線への適用を可能とした伸縮性を有する配線基板およびその製造方法。
特開2023-181815	伸縮性を有する配線基板及びその製造方法	高い伸縮性と導電性とを兼ね備えた導電性薄膜によって、微細な配線への適用を可能とした伸縮性を有する配線基板およびその製造方法。
特開2023-182937	反応管理装置及びプログラム	例えば視聴者の反応がリアルタイムに把握しにくい遠隔地からの観覧であっても、視聴者の反応を撮影スタジオに反映することができる反応管理装置およびプログラム。
特開2023-182982	映像伝送システム、映像伝送装置、及び映像伝送プログラム	映像伝送時のソフトウェアによる圧縮伸張処理における単一リソース規模の低減と低遅延化を可能とする、映像分割による圧縮伸張処理の分散処理化と分散処理の自動化とを提供する映像伝送システム、映像伝送装置、および映像伝送プログラム。
特開2023-183538	撮像素子及び撮像素子の製造方法	結晶セレン膜を光電変換部に用いた従来の撮像素子の応用範囲の拡大が困難である課題に対して、応用範囲を拡大することができる撮像素子および撮像素子の製造方法。
特開2023-183838	透光性基板、接合型固体撮像素子、および透光性基板の製造方法	製造上の手間やコストを抑えた接合型固体撮像素子が備える透光性基板、接合型固体撮像素子、および透光性基板の製造方法。
特開2024-63	アクセント位置推定装置、アクセント位置学習装置及びそれらのプログラム	単語の意味を表す単語分散表現ベクトルに加え、単独発話時のアクセント位置を表す単語素性ベクトルが反映された深層学習モデルを用いることで、あるいはそれを学習に使用することで、アクセント位置の推定精度を向上させるアクセント位置推定装置、アクセント位置学習装置およびそれらのプログラム。
特開2024-443	映像符号化装置及びプログラム	HDR (High Dynamic Range: 高ダイナミックレンジ) / 広色域映像のような、鮮やかな色彩を含む映像を符号化する際の主観品質を向上させることが可能な映像符号化装置およびプログラム。

公開番号	発明考案の名称	技術概要
特開2024-1432	IP番組切替装置及びIP番組切替プログラム	ロケーションの異なる拠点で制作されたライブ番組素材を、例えばクラウド上でも指定時刻によりフレーム精度で切り替えることが可能となり、放送のような高精度な番組切替機能を持ったリニア配信サービスをクラウド上でも運用可能とするためのIP (Internet Protocol) 番組切替装置およびIP番組切替プログラム。
特開2024-2255	撮像装置	3板撮像方式をはじめとする多板撮像方式の撮像装置であって、簡便な方法で、多方向の合焦状態検出と、位相差検出画素による情報欠落が少ない高画質な画像取得を両立することができる撮像装置。
特開2024-2467	固体撮像素子の製造方法	結晶セレン膜の緻密性を維持するとともに、暗電流を低減させることが可能な固体撮像素子の製造方法。
特開2024-2497	薄膜トランジスタを備えた撮像素子及びその製造方法	垂直色分離型の撮像素子において、画素開口率を層間で変化させることなく、かつ画素開口率を低下させることなく各層にて適切な大きさの蓄積容量を設けた撮像素子およびその製造方法。
特開2024-2509	ストリーム送出装置及びストリーム送出プログラム	緊急時の文字スーパー上乗せなどの送出コンテンツの即時加工や、送出寸前でもコンテンツの部分または全部の差し替えがワークフローを変更することなく実現できる、プリエンコード済みストリームの送出装置および送出プログラム。
特開2024-3218	復号装置及びプログラム	画像を構成するブロック単位で復号を行う復号装置であって、符号化ストリームを復号することで出力された復号対象ブロックの変換係数に対して、変換係数に含まれる最終有意係数座標 (Last Position) の符号化の効率を改善する復号装置およびプログラム。
特開2024-3343	送信装置及び受信装置	地上デジタル放送の高度化方式において、受信装置の回路規模の増大を抑制しながら、TxID (Transmitter Identification: 送信局識別信号) を適切に送信することを可能とする送信装置および受信装置。
特開2024-3368	ルータ	IP (Internet Protocol) ネットワークにおいて、マルチキャストトラフィックの無駄な転送を自動的に防ぐ機能を、少ないメモリ量で実現するルータ。
特開2024-3674	IP/TS変換装置及びプログラム	IP (Internet Protocol) 回線を用いた伝送システムにおいて、ユーザーの設定に従い、IPデータとTS (Transport Stream) との間の変換処理の際に、IPフローごとに異なるデータ量、許容できる遅延時間および必要となる冗長度が異なることに対して、IPフローごとの伝送を制御することができるIP/TS変換装置およびプログラム。
特開2024-3688	撮像素子及び撮像装置	撮像領域ごとに局所的な高フレームレート撮影若しくは高解像度撮影による撮像を実現し、かつ、撮影モードの異なる領域の境界部分に生じうるアーティファクトを軽減し、画面全体の画質を向上させる撮像素子および撮像装置。
特開2024-3786	送信装置及び受信装置	地上デジタル放送の高度化方式において、受信装置の処理負荷や消費電力の増大を抑制しながら、TxID (Transmitter Identification: 送信局識別信号) を適切に受信することを可能とする送信装置および受信装置。
特開2024-4146	表示装置、およびプログラム	光線の方向を制御するためのアレー素子における素子に対応して投影される画像を用いて、インテグラル方式 (光線の方向を制御して実際に物体が存在する場合と同じ光線の状態を創出することにより、光線を発生させる表示面とは異なる奥行き位置に立体像を表示する方式) により立体像を表示する表示装置であって、表示性能を容易に変更できる表示装置およびプログラム。
特開2024-4407	送信装置及び受信装置	地上デジタル放送の高度化方式において、受信装置の回路規模の増大を抑制しながら、TxID (Transmitter Identification: 送信局識別信号) を適切に送信することを可能とする送信装置および受信装置。
特開2024-4564	信号処理装置及び信号処理方法	映像信号と、映像信号に所定の変化があったことを示す差分信号とを、デジタル値として容易に取得できる信号処理装置および信号処理方法。
特開2024-4565	信号処理装置及び信号処理方法	映像信号と、映像信号に所定の変化があったことを示す差分信号とを、デジタル値として容易に取得できる信号処理装置および信号処理方法。

公開番号	発明考案の名称	技術概要
特開2024-5644	合焦制御装置、表示装置、および画像伝送装置	画像を観視する際、注視領域では合焦した画像を観視者が疲労しないように提示することが可能であり、注視点のない領域では情報量を抑制し、伝送データ量を削減することが可能な合焦制御装置、表示装置、および画像伝送装置。
特開2024-7248	配信サーバ、受信装置、及びプログラム	低遅延セグメント形式のストリーミングコンテンツを、低遅延セグメント形式に非対応な受信装置でも再生できるようにする配信サーバ、受信装置、およびプログラム。
特開2024-7266	ネットワークノード及びプログラム	送信装置から受信したオブジェクトデータをフィルタ処理して受信装置に送信するネットワークノードにおいて、受信装置への3次元オブジェクトデータの伝送効率を向上させることが可能なネットワークノードおよびプログラム。
特開2024-7629	映像符号化装置、プリ処理装置、及び映像符号化方法	ニューラルネットワークを用いて符号化効率の向上を図る映像符号化装置、プリ処理装置、および映像符号化方法。
特開2024-7635	映像処理装置、アップスケーリング方法、及び映像処理システム	ニューラルネットワークを用いてアップスケーリングすることで、符号化効率を向上させるようにした映像処理装置、アップスケーリング方法、および映像処理システム。
特開2024-7667	端末装置、配信装置及びプログラム	横長動画を縦長の表示画面で好適に視聴することが可能な端末装置、配信装置およびプログラム。
特開2024-8262	字幕データ生成装置及び字幕データ生成プログラム	字幕データの標準に準拠しつつ、個人に合わせて内容を変更できる字幕データ生成装置および字幕データ生成プログラム。
特開2024-10859	変換装置およびプログラム	映像を単語列に変換する変換装置において、正例（正解の単語列）だけでなく、負例（不正解の単語列）による機械学習も可能とする変換装置およびプログラム。
特開2024-11020	パッケージ、半導体装置及び半導体装置の製造方法	上面にパッド電極を有する半導体チップが、内部と外部とを接続するリード部を有するパッケージに包装されている半導体装置であって、ワイヤボンディングの領域を確保しながら小型化を図ることができるパッケージ、半導体装置および半導体装置の製造方法。
特開2024-11063	送信装置、受信装置及びプログラム	地上デジタル放送の高度化方式において、TMCC（Transmission and Multiplexing Configuration Control：伝送制御情報）情報の伝送特性を改善することを可能とする送信装置、受信装置およびプログラム。
特開2024-11160	姿勢推定装置、学習装置及びプログラム	動画に含まれる被写体の姿勢を推定する際に、2次元骨格情報からモーションジッターを除去し、被写体の自然な動きを実現する姿勢推定装置、学習装置およびプログラム。
特開2024-11713	表情アニメーションデータ生成装置およびそのプログラム	顔の表情をより正確に表現するとともに、CGソフトウェア上で動作や変更が可能なアニメーションデータを生成する表情アニメーションデータ生成装置およびそのプログラム。
特開2024-11770	撮像素子、撮像装置および撮像素子の製造方法	撮像素子が大面積で画素数が大幅に増加した場合であっても、画素信号読出し時における1ライン分に要する選択時間を短縮することができ、動画に必要な枚数の画像を取得することが可能な撮像素子、撮像装置および撮像素子の製造方法。
特開2024-13279	視聴予測装置	対象ユーザーの現在までの視聴傾向から、5年先、10年先といった長いスパンでの対象ユーザーの視聴傾向を予測することを可能とする視聴予測装置。
特開2024-13280	感情判定装置およびモデル学習装置、ならびに、それらのプログラム	画像には映っていない顔の情報を補助情報として利用することで、画像に映った人物の感情を精度よく判定することが可能な感情判定装置およびモデル学習装置、ならびにそれらのプログラム。

公開番号	発明考案の名称	技術概要
特開2024-13638	透光性基板、および接合型固体撮像素子	従来の接合型固体撮像素子は、製造時に多くの加熱工程や冷却工程にさらされてさまざまな応力が加わることでわずかに湾曲しているCMOS回路基板上に形成された第1非晶質セレン膜と、確立された研磨工程により平坦性の高い透光性基板上に形成された第2非晶質セレン膜とが接合されていたため、これらの非晶質セレン膜が密着し難く歩留まりが低下する問題があったことに対し、その歩留まりを向上させる透光性基板および接合型固体撮像素子。
特開2024-13728	色分布表示装置およびそのプログラム	3次元の色空間で表現された色域立体の色域データを2次元で表示する色域リング上に、色域が対応する画像データの色分布を重畳して表示する色分布表示装置であって、色域リングにおいて従来よりも幅の広いリング内で彩度の表現をより明確に視覚化して、画像の色分布を表現することができる色分布表示装置およびそのプログラム。
特開2024-15616	視聴者装置、システム、およびプログラム	現実空間を映した映像内において、ユーザーの運動の結果に基づく位置にアバターを表示させることのできる視聴者装置、システム、およびプログラム。
特開2024-15897	送信装置およびプログラム	デジタル信号を所定の変調方式で変調した変調波信号を、所定の伝送路を介して送信する送信装置であって、ハードウェア規模を低減しつつ、ひずみ補償機能を実現することができる送信装置およびプログラム。
特開2024-15899	送信システムおよびプログラム	デジタル信号が所定的方式で変調された変調波信号を生成する送信装置と、所定の伝送路による伝送路ひずみを補償するための信号成分を変調波信号に付加する信号処理装置とを備える送信システムであって、ハードウェア規模を低減しつつ、ひずみ補償機能を実現することができる送信システムおよびプログラム。
特開2024-16541	撮像装置およびプログラム	単体の撮像装置で遅延を生じさせることなく、被写体の動作速度に応じてフレームレートを切り替えることのできる撮像装置およびプログラム。
特開2024-16542	撮像装置	垂直同期信号などの同期信号に同期して駆動される固体撮像素子の画素の設計(画素の配置等)を変更せずに、イベント検出型画素を有する撮像素子を用いて、動き検出の精度を高めることができる撮像装置。
特開2024-16955	符号化装置、ストリーム合成装置、復号装置、およびプログラム	符号処理を行う装置側に大きな負荷をかけずに、復号装置ごとに最適なビットストリームを生成することのできる符号化装置、ストリーム合成装置、復号装置、およびプログラム。
特開2024-17080	バックプレーン	ホログラム像の品質低下を抑制できるバックプレーン。
特開2024-17570	磁壁移動型空間光変調器	簡易なプロセスで、駆動電流を大きくせずに、磁壁移動型空間光変調素子の開口率を大きくすることが可能な磁壁移動型空間光変調器。
特開2024-17904	触覚提示装置、自己速度推定装置及びプログラム	カメラの位置を視聴者の位置としてカメラが移動することで撮影された一人称視点映像を視聴する際に、没入感向上に寄与する触覚刺激を提示するための情報を生成する触覚提示装置およびプログラム、ならびに一人称視点映像から精度の高い自己速度を推定できる自己速度推定装置およびプログラム。
特開2024-17905	触覚提示装置及びプログラム	視聴者が映像を視聴する際に、没入感向上に寄与する触覚刺激を提示するための情報を生成する触覚提示装置およびプログラム。
特開2024-18058	機械学習システム	ユーザーの匿名性を保証し、プライバシーを保護できる機械学習システム。
特開2024-18454	符号化撮像装置	低解像度符号化画像から高解像度画像を再構成する際に、マスクされた画素領域に生じるノイズを原因とした、ダイナミックレンジの低下に伴う、画像全体の輝度レベルの低下や持ち上がりを抑制し、画質の高い高解像度画像を取得できる符号化撮像装置。
特開2024-19542	復号装置、プログラム、及び復号方法	複数の色成分により構成される原画像を分割して得たブロックに対する復号処理を行う復号装置であって、色差残差スケールングを適切に適用することで符号化効率を改善する復号装置、プログラム、および復号方法。

公開番号	発明考案の名称	技術概要
特開2024-19856	衛星追尾装置	人工衛星からの信号を受信するためのアンテナの方向調整を自動的に行って、精度よく追尾する衛星追尾装置。
特開2024-19868	衛星放送受信用の周波数変換装置、及び電力測定システム	イメージ干渉の影響を低減する衛星放送受信用の周波数変換装置、および周波数変換装置を備え、その出力から得られる各チャンネルの中間周波数信号の受信電力を測定する電力測定システム。
特開2024-20013	検査装置、検査システム、及びプログラム	有機EL素子について、重大な画質破綻を引き起こすシュリンク（有機EL素子の電極金属または有機膜が水分子と反応した部分）は非発光領域となり、それが素子端部および基板上の微小な異物粒子付近で時間的に拡大していく、このような素子端部に生じた劣化のことでとダークスポット（素子上の微小な異物粒子付近に生じた劣化のことで）を区別して、高い精度で劣化部分を解析することを可能にする検査装置、検査システム、およびプログラム。
特開2024-20588	符号化装置、復号装置、及びプログラム	画像を構成する輝度信号および色差信号を分割して生成した輝度ブロックおよび色差ブロックを符号化／復号化する符号化装置／復号装置であって、色差信号のイントラ予測モードとしてDM（Derived Mode：対応する輝度ブロックに適用された予測モード）を選択した場合の符号化効率を改善する符号化装置、復号装置、およびプログラム。
特開2024-21255	点群表現学習用オートエンコーダー装置、点群表現エンコーダー装置、およびプログラム	点群（Point Cloud）の相対位置の情報のみに基づいて、それらの点群の絶対位置の情報を含んだ点群表現を推定することが可能な、点群の絶対座標を汎化することができる点群表現学習用オートエンコーダー装置、点群表現エンコーダー装置、およびプログラム。
特開2024-21715	領域検出装置およびプログラム	ニューラルネットワークなどの機械学習モデルを用いて画像内における特定の特徴を有する領域を検出ことができ、検出される領域が長方形に限定されることのない領域検出装置およびプログラム。
特開2024-21896	良視点画像選択装置、及びコンピュータを良視点画像選択装置として機能させるためのプログラム	複数のオブジェクトが存在する複雑なシーンにおいて、最適な視点を検出し、その検出した最適な視点から見た画像を生成する良視点画像選択装置、およびコンピュータを良視点画像選択装置として機能させるためのプログラム。
特開2024-22374	符号化装置、復号装置、及びプログラム	映像を所要ビットレートで符号化する際に、計算処理を軽量化しつつ、画質および符号化効率を向上させることが可能な符号化装置、復号装置、およびプログラム。
特開2024-23163	音声信号処理装置およびプログラム	複数の音声オブジェクトから構成される番組音声の全体的な印象の悪化を抑制しつつ、特定の音声オブジェクトの聞き取りやすさの向上を図ることができる音声信号処理装置およびプログラム。
特開2024-25854	検索装置、固有表現抽出装置、およびプログラム	対象文と意味的に類似した文以外に、固有表現抽出（文章中の人名や地名や組織名などの固有表現を抽出して、「人名」、「地名」、「組織名」などのクラスに分類する技術）の精度の向上に寄与し得る文を外部データから取得することのできる検索装置あるいはそのプログラム、およびそのような検索装置を利用した固有表現抽出装置あるいはそのプログラム。

# 学会発表論文一覧 (2023年7月～2023年12月)

論文名	発表者	誌名	巻号
Development of a Measurement System for the Radiation Characteristics of the Human Voice	木下光太郎, 杉本岳大	Acoustical Science and Technology	Vol.44, No.4, pp.344-347
[Invited Paper]End-to-end Verification of the Advanced Broadcasting System	朝倉慎悟, 永田裕靖, 兜森 椋, 大西正芳, 神原浩平, 大槻一博, 土田健一 (NHK財団)	ITE Transactions on Media Technology and Applications	Vol.11, No.3, pp.113-122
Synthesis of Sound Field from Moving Complex Sources with Arbitrary Trajectories by Linear and Spherical Loudspeaker Arrays	佐々木陽, 松井健太郎, 中山靖茂	Journal of acoustical society of America	Vol.154, No.1, pp.571-588
マスク着用に対応した顔認識手法	河合吉彦, 望月貴裕, 苗村昌秀 (NHK財団)	映像情報メディア学会誌	Vol.77, No.4, pp.537-545
Effects of Energy-Level Alignment on Operating Voltages of Blue Organic Light-Emitting Diodes	佐々木翼, 大野 拓, 清水貴央, 深川弘彦	Advanced Materials Interfaces	Vol.10, No.19, 2201925
Angular Resolution of Radiation Characteristics Required to Reproduce Uttered Speech in all Three-dimensional Directions	杉本岳大, 木下光太郎	Acoustical Science and Technology	Vol.44, No.5, pp.360-370
Pixel-Parallel Three-Layer Stacked CMOS Image Sensors Using Double-Sided Hybrid Bonding of SOI Wafers	後藤正英, 本田悠葵, 難波正和, 井口義則 (NHK財団), 更屋拓哉 (東京大学), 小林正治 (東京大学), 日暮栄治 (東北大学), 年吉 洋 (東京大学), 平本俊郎 (東京大学)	IEEE Transactions on Electron Devices	Vol.70, No.9, pp.4705-4711
地上テレビジョン放送高度化方式に対応するMFN放送波中継用等化判定器	川島祥吾, 宮坂宏明, 朝倉慎悟, 竹内知明, 岡野正寛	映像情報メディア学会誌	Vol.77, No.5, pp.670-680
予測モデルを用いたサッカー中継用AIロボットカメラ装置による試合会場での自動撮影実験	高橋康太, 三須俊枝, 久富健介	映像情報メディア学会誌	Vol.77, No.5, pp.690-693
Spectrally Narrow Band-edge Photoluminescence from AgInS <sub>2</sub> -based Core/Shell Quantum Dots for Electroluminescence Applications	上松太郎 (大阪大学), 泉竜之介 (大阪大学), 菅野祥生 (大阪大学), 菅野 陸 (大阪大学), 平野達也 (大阪大学), 本村玄一, 鳥本 司 (名古屋大学), 桑畑 進 (大阪大学)	Faraday Discussions [Early View]	DOI:10.1039/D3FD00142C, <https://doi.org/10.1039/D3FD00142C>, (First published: 02 Sep. 2023)
Field Experiment with FPU for Mobile Relay Programs with Auto-switching of Total Modulation Orders	牧野仁宣, 佐藤貴弘, 伊藤史人, 中川孝之, 居相直彦	ITE Transactions on Media Technology and Applications	Vol.11, No.4, pp.164-175
車室サイズの聴取空間における22ch 音響の評価	木村里帆 (山梨大学), 小澤賢司 (山梨大学), 丹野慶太 (アルプスアルパイン), 伊勢友彦 (アルプスアルパイン), 北島 周, 久保弘樹, 大出訓史, 大久保洋幸 (NHK財団)	日本音響学会誌	Vol.79, No.10, pp.496-501
Program-Production System Consisting of Multiple NMOS-Compliant Systems	小山智史, 佐藤俊喜 (日本電気), 打田良介 (日本電気), 菅原正幸 (日本電気), 倉掛卓也	SMPTE Motion Imaging Journal	Vol.132, No.9, pp.26-31
Ferroelectric liquid Crystal Array Driven by a Two-layer Electrode with a 1×1 μm Pixel Pitch for Light Modulation in Electro-holography	麻生慎太郎, 山本貴翔 (東北大学), 青島賢一, 東田 諒, 船橋信彦, 柴崎純一, 石鍋隆宏 (東北大学), 柴田陽生 (長岡技術科学大学), 藤掛英夫 (東北大学), 町田賢司	Journal of the Society for Information Display[Early View]	DOI:10.1002/jsid.1260, <https://doi.org/10.1002/jsid.1260>, (First published: 05 Oct. 2023)
Bright Red-color Luminescence from Ag-In-Ga-S-based Quantum Dots with the Introduction of Copper	上松太郎 (大阪大学), Navapat Krobkrong (大阪大学), 浅井健一郎 (大阪大学), 本村玄一, 藤崎好英, 鳥本 司 (名古屋大学), 桑畑 進 (大阪大学)	Bulletin of the Chemical Society of Japan	Vol.96, No.11, pp.1274-1282
Direct Patterned Printing Method of Top-Contact ITO Electrodes for Solution-Processed Metal Oxide Thin-film Transistors	宮川幹司, 辻 博史, 中田 充	ACS Applied Electronic Materials	Vol.5, No.11, pp.6056-6062
One-pot Synthesis of Ag-In-Ga-S Nanocrystals Embedded in a Ga <sub>2</sub> O <sub>3</sub> Matrix and Enhancement of Band-edge Emission by Na <sup>+</sup> Doping	都澤 諒 (名古屋大学), 宮前千恵 (名古屋大学), 秋吉一孝 (名古屋大学), 亀山達矢 (名古屋大学), 山本剛久 (名古屋大学), 本村玄一, 藤崎好英, 上松太郎 (大阪大学), 桑畑 進 (大阪大学), 鳥本 司 (名古屋大学)	Nanoscale Advances[Early View]	DOI: 10.1039/D3NA00755C, <https://doi.org/10.1039/D3NA00755C>, (First published: 14 Nov. 2023)



## 学会発表論文一覧 (2023年7月～2023年12月)

論文名	発表者	誌名	巻号
Demonstration Experiment of a System that Helps Users Select Educational Content Using Terms in Basic/advanced Relationships	藤井翔子, 阿部晋矢, 藤津 智, 松村欣司, 藤沢 寛	Journal of Information Processing	Vol.31, pp.766-774
Capturing Videos at 60 Frames per Second Using Incoherent Digital Holography	室井哲彦, 信川輝吉, 片野祐太郎, 萩原 啓	Optics Continuum	Vol.2, No.11, pp.2409-2420
Machine Learning-Based Compensation Methods for Weight Matrices of SVD-MIMO	牧野仁宣, 中川孝之, 居相直彦	IEICE Transactions on Communications	Vol.E106-B, No.12, pp.1441-1454
Posner課題を用いた微風による注意誘導の検討	西澤昌宏, 原澤賢充, 半田拓也, 小峯一晃	日本バーチャルリアリティ学会論文誌	Vol.28, No.4, pp.371-380

# 研究会・年次大会等発表一覧 (2023年11月～2024年2月)

題目	発表者	発表先/誌名	資料番号	発表年月日
ドナー・アクセプター界面のエネルギー解析による低電圧駆動OLEDの実現	佐藤千夏 (東京理科大学), 岡田拓也, 大野 拓, 佐々木翼, 清水貴央, 深川弘彦	2023年第37回有機EL討論会例会予稿集	S8-5, pp.39-40	2023.11.16 ～ 17
Unlocking the Full Potential of Electron Acceptor Molecules for Efficient and Stable Hole Injection into Organic Semiconductors	大野 拓, 岡田拓也, 佐々木翼, 清水貴央, 深川弘彦	Materials Research Society Symposium 2023 [MRS Fall Meeting 2023]	EL16.07.03	2023.11.26 ～ 12.1
Artificial-Intelligence Robotic Camera System for Live Soccer Broadcasts	高橋康太, 三須俊枝, 久富健介	20th ACM SIGGRAPH European Conference on Visual Media Production (CVMP 2023) Programme Booklet	p.26	2023.11.30 ～ 12.1
[Invited] Common Metals as Electrodes for Organic Light-Emitting Devices	大久 哲, 大野 拓, 宮川幹司, 清水貴央	4th International Conference on Optics, Photonics, and Lasers (OPL-2023)	[Session: Invited Talks]	2023.12.4 ～ 7
Fabrication of Quantum Dot Light-Emitting Diode Display with Multinary Compound Semiconductor Quantum Dots	本村玄一, 大久 哲, 上松太郎 (大阪大学), 桑畑 進 (大阪大学), 亀山達矢 (名古屋大学), 鳥本 司 (名古屋大学), 藤崎好英	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, FMC1-3, pp.319-322	2023.12.4 ～ 9
High-Resolution Aerial Image by Placing a Ball Lens at a Virtually Conjugate Position in AIRR Optical System	滝山和晃 (宇都宮大学), 陶山史朗 (宇都宮大学), 正岡顕一郎, 山本裕紹 (宇都宮大学)	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, FMC6-1, pp.361-364	2023.12.4 ～ 9
Gamut Rings Color Scope for Signal Monitoring	正岡顕一郎	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, FMC11-3L, pp.407-410	2023.12.4 ～ 9
Pixel-Density Enhanced Integral Three-Dimensional Display by Time-Division Multiplexing of Two-Dimensional Image	渡邊隼人, 岡市直人, 加納正規, 佐々木久幸, 洗井 淳	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, 3D4/3DSA4-3, pp.660-663	2023.12.4 ～ 9
Aperture Control for Increasing Hologram Signal Amplitude in Incoherent Digital Holography	室井哲彦, 信川輝吉, 片野祐太郎, 萩原 啓	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, 3D8/3DSA8-2, pp.684-685	2023.12.4 ～ 9
Lateral Electric Field Driving of High-Resolution Liquid Crystal Devices for Electronic Holography with Wide Viewing-Zone Angle	栃木硯人 (東北大学), 中谷誠和 (東北大学), 青島賢一, 川那真弓, 山口祐太, 町田賢司, 船橋信彦, 藤掛英夫 (東北大学)	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, 3Dp1-8L, pp.715-718	2023.12.4 ～ 9
[Invited] "Otogi Box": A Cubic Haptic Interface to Tell a Story in Your Hands	東真希子, 半田拓也, 小峯一晃	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, VHF1-1, pp.719-722	2023.12.4 ～ 9
[Invited] Required Specifications on Spatial Aspects for Ideal HMD Estimated on Head-Centered Coordinate System	原澤賢充, 宮下山斗, 小峯一晃	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, VHF4-1L, pp.748-750	2023.12.4 ～ 9
Thermofiber: a Non-Contact, Head-Mounted Spatial Thermal Display Using Visible Light Heating by Optical Fibers	西澤昌宏, 原澤賢充, 半田拓也, 小峯一晃	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, VHF4-3L, pp.754-757	2023.12.4 ～ 9
A Wide-Viewing and High-Resolution Head-Mounted Display Using Multiple Display Panels	水谷沙耶, 前田恭孝, 半田拓也, 久富健介, 小出大一	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, PRJ7-4L, pp.838-841	2023.12.4 ～ 9
[Invited] Development of Highly Stretchable 3-axis Deformable LED Displays	宮川幹司, 辻 博史, 武井達哉, 山本敏裕 (NHK財団), 藤崎好英, 中田 充	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, FLX3-3L, pp.1129-1131	2023.12.4 ～ 9
Flexible Bezel-less Micro-LED Display Driven from Substrate Back Side Using Through-plastic Vias	辻 博史, 武井達哉, 宮川幹司, 山本敏裕 (NHK財団), 本村玄一, 藤崎好英, 中田 充	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, FLX3-4L, pp.1132-1134	2023.12.4 ～ 9
[Invited] Incoherent Digital Holography Aiming to Capture Three-Dimensional Video	室井哲彦, 信川輝吉, 片野祐太郎, 萩原 啓	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, INP3/LCT6-1, pp.1209-1212	2023.12.4 ～ 9
Pixel-wise Phase-shifting Incoherent Digital Holography with a Liquid Crystal Lens	信川輝吉, 臼井真広, 片野祐太郎, 後藤正英, 萩原 啓, 室井哲彦	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, INP3/LCT6-2, pp.1213-1214	2023.12.4 ～ 9

# 研究会・年次大会等発表一覧 (2023年11月～2024年2月)

題目	発表者	発表先/誌名	資料番号	発表年月日
[Invited] Holographic Image Reconstruction Using a Magneto-Optical Spatial Light Modulator with its Pixel Pitch of 1 $\mu$ m and its Layout of 10k $\times$ 5k	青島賢一, 船橋信彦, 東田 諒, 川那真弓, 山口祐太, 町田賢司	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, 3DSA1/3D1-2, pp.1253-1256	2023.12.4～9
Improved Eye-Tracking Function by Combining Inertial Sensor Information for Integral 3D Display	佐々木久幸, 加納正規, 岡市直人, 渡邊隼人, 三科智之	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, 3DSA3/3D3-2, pp.1266-1269	2023.12.4～9
Point Light Source Array Based Integral 3D Display Using Eye-Tracking and Time-Division Technologies	岡市直人, 渡邊隼人, 加納正規, 佐々木久幸, 洗井 淳	Proceedings of the 30th International Display Workshops (IDW '23)	Vol.30, 3DSA5/3D5-2, pp.1276-1279	2023.12.4～9
Flexible Integrated Circuits Developed by Transferring FDSOI CMOS FETs on Plastic Substrate for CMOS Image Sensors	後藤正英, 為村成亨, 佐藤弘人	25th IEEE Electronics Packaging Technology Conference (EPTC 2023)	IEEE, pp.237-240	2023.12.5～8
短時間視聴時代の映像解析技術 ～映像自動要約・サムネイル自動抽出～	望月貴裕, 前澤桃子	ビジョン技術の実利用ワークショップ ViEW2023 [精密工学会画像応用技術専門委員会主催]	OS4-K1, pp.366-373	2023.12.7～8
映像における顔の見えない人物のラベル推定手法の検討	藤森真綱, 望月貴裕	ビジョン技術の実利用ワークショップ ViEW2023 [精密工学会画像応用技術専門委員会主催]	OS4-O1, pp.374-378	2023.12.7～8
高臨場感メディアにおける高速度イメージング	北村和也	高速度イメージングとフォトニクスに関する総合シンポジウム2023 [JCHSIP2023]	14-1	2023.12.13～15
画面内の領域ごとに解像度・フレームレート・シャッタースピードを制御可能なイメージセンサー	富岡宏平, 菊地幸大, 薄井武順, 北村和也, 川人祥二 (静岡大学)	高速度イメージングとフォトニクスに関する総合シンポジウム2023 [JCHSIP2023]	14-2	2023.12.13～15
マイケルソン干渉計を用いたインコヒーレントデジタルホログラフィによる3次元動画画像の取得	室井哲彦, 信川輝吉, 片野祐太郎, 萩原 啓	高速度イメージングとフォトニクスに関する総合シンポジウム2023 [JCHSIP2023]	14-4	2023.12.13～15
ゴルフ中継用ロボティクス制作システムの検討	加藤大一郎 (NHK財団), 三ツ峰秀樹 (NHK財団), 久富健介, 三須俊枝	第24回計測自動制御学会システムインテグレーション部門講演会 (SI2023)	1G1-05, pp.1171-1172	2023.12.14～16
イマーシブメディアの実現を目指したデバイス技術	清水貴央	日本学術振興会光電相互変換第125委員会「本委員会第266回研究会」研究会資料	pp.120-129	2023.12.15
顔遮蔽検出を用いた手話CG顔モーション修正手法の提案	箱崎浩平, 村上智哉, 高橋正樹, 金子浩之	2023年映像情報メディア学会冬季大会講演予稿集	11B-3	2023.12.26～27
テレビ番組への興味に基づくロボットの発話文生成手法の検討	萩尾勇太, 上村真利奈, 大亦寿之, 奥田 誠	2023年映像情報メディア学会冬季大会講演予稿集	11B-7	2023.12.26～27
手話映像と日本語対訳を用いたマルチモーダル手話認識	梶山岳士, 宮崎太郎, 金子浩之	2023年映像情報メディア学会冬季大会講演予稿集	11B-8	2023.12.26～27
文字起こしのためのオープンソース汎用認識モデルの活用	奥 貴裕, 三島 剛 (NHK財団), 菅野竜雅, 佐久間旭, 佐藤裕明, 熊野 正, 河合吉彦	2023年映像情報メディア学会冬季大会講演予稿集	13B-1	2023.12.26～27
8K-ROIスイッチャー制作システムの開発	薄井武順, 三須俊枝, 高橋正樹, 遠藤 伶, 北村和也	2023年映像情報メディア学会冬季大会講演予稿集	13C-1	2023.12.26～27
市販産業用カメラを用いた五角柱型30K360度カメラの設計	林田哲哉, 米内 淳, 原 一宏, 大川裕司, 瀧口吉郎	2023年映像情報メディア学会冬季大会講演予稿集	13C-2	2023.12.26～27
奥行画像を用いたスティッチング技術の検討	原 一宏, 林田哲哉, 瀧口吉郎	2023年映像情報メディア学会冬季大会講演予稿集	13C-3	2023.12.26～27
曲面ディスプレイに表示する広視野高精細コンテンツの制作手法の検討	原 一宏, 相澤明香里, 高橋廉太郎, 菊池哲男, 吉田孝光	2023年映像情報メディア学会冬季大会講演予稿集	13C-4	2023.12.26～27
ユーザ主体のパーソナルデータ管理モデルによるSNS投稿履歴情報に基づいた番組推薦システムの試作	上野幹大, 藤津 智	2023年映像情報メディア学会冬季大会講演予稿集	13C-5	2023.12.26～27
簡易受信を可能とする12GHz帯衛星放送システムの提案 ～超低C/N伝送方式と小型受信アンテナの研究開発～	亀井 雅, 長坂正史, 小島政明, 小泉雄貴, 秋山良太	2023年映像情報メディア学会冬季大会講演予稿集	13C-6	2023.12.26～27
有線IP回線からのパケット補完機能を備えた高度化STL/TTLシステムのフィールド実験	梅内哲也, 島崎智拓, 永田裕靖, 居相直彦, 中川孝之	2023年映像情報メディア学会冬季大会講演予稿集	13C-7	2023.12.26～27

# 研究会・年次大会等発表一覧 (2023年11月～2024年2月)

題目	発表者	発表先／誌名	資料番号	発表年月日
地上テレビジョン放送高度化方式における固定受信向けサブフレームの伝送容量の増加に関する検討	宮坂宏明, 竹内知明, 岡野正寛	2023年映像情報メディア学会冬季大会講演予稿集	13C-8	2023.12.26～27
地上放送高度化方式の機能検証 ～物理層分割モードチャンネルボンディングの野外実験～	朝倉慎悟, 佐藤明彦, 宮坂宏明, 神原浩平, 岡野正寛	2023年映像情報メディア学会冬季大会講演予稿集	13C-9	2023.12.26～27
野球中継映像のみを用いた投球コース推定	島野雄貴, 栗野裕也, 高橋正樹, 金子浩之	2023年映像情報メディア学会冬季大会講演予稿集	21B-4	2023.12.26～27
効率的な3D映像制作のための良視点自動推定技術	界 瑛宏, 澤島康仁, 宮下山斗, 原澤賢充, 小峯一晃	2023年映像情報メディア学会冬季大会講演予稿集	22B-7	2023.12.26～27
フレキシブルTFTイメージセンサの作製プロセスの開発	今村弘毅, 宮川和典, 為村成亨, 佐藤弘人	2023年映像情報メディア学会冬季大会講演予稿集	23C-1	2023.12.26～27
シーン適応型撮像システムにおける動き領域の判定精度評価に関する一検討	本地 瑛, 菊地幸大, 富岡宏平, 薄井武順, 北村和也	2023年映像情報メディア学会冬季大会講演予稿集	22C-4	2023.12.26～27
乾電池一本で発光する緑色・青色OLEDの開発	佐藤千夏 (東京理科大学), 岡田拓也, 大野 拓, 佐々木翼, 清水貴央, 深川弘彦	2023年映像情報メディア学会冬季大会講演予稿集	23C-5	2023.12.26～27
多元系半導体量子ドットを使った高色純度発光素子の開発	本村玄一, 上松太郎 (大阪大学), 亀山達矢 (名古屋大学)	2023年映像情報メディア学会冬季大会講演予稿集	企画4-6	2023.12.26～27
Hardware Implementation of Low-Latency Image Denoising with Wavelet Shrinkage and Noise2Noise-Based Learning for Ultra-High Definition Cameras	菊地幸大, 富岡宏平, 本地 瑛, 薄井武順, 平川恵悟 (デイトン大学), 北村和也	42nd IEEE International Conference on Consumer Electronics (ICCE 2024)	[Session CT02-1: Audio/Video Systems and Signal Processing I (AVS)]	2024.1.5～8
[招待講演] アクティナビジョン：高密度な光線再生によるフルバラックス3次元映像表示技術	渡邊隼人	レーザー学会学術講演会第44回年次大会講演論文集	S12-17a-VIII-04	2024.1.16～19
横電界駆動を用いた電子ホログラフィ用超高解像度液晶デバイスの設計	栃木碩人 (東北大学), 中谷誠和 (東北大学), 青島賢一, 川那真弓, 山口祐太, 町田賢司, 船橋信彦, 藤掛英夫 (東北大学)	映像情報メディア学会技術報告	Vol.48, No.1, IDY2024-9, pp.33-36	2024.1.25～26
動画ホログラフィのための高解像度駆動用強誘電性液晶素子の低電圧化と高速化	小林弘輝 (東北大学), 栃木碩人 (東北大学), 中谷誠和 (東北大学), 青島賢一, 川那真弓, 山口祐太, 町田賢司, 船橋信彦, 藤掛英夫 (東北大学)	映像情報メディア学会技術報告	Vol.48, No.1, IDY2024-11, pp.41-44	2024.1.25～26
Demonstration of Optical Beam Manipulation with Optical Phased Array Using EO Polymer and Si <sub>3</sub> N <sub>4</sub> Dual-Layer Waveguides	宮本裕司, 三浦雅人, 柴崎純一, 難波正和, 町田賢司, 上田里永子 (情報通信研究機構), 梶 貴博 (情報通信研究機構), 山田俊樹 (情報通信研究機構), 大友 明 (情報通信研究機構), 平野芳邦	SPIE Photonics West 2024 (SPIE OPTO)	Paper 12889-68	2024.1.27～2.1
[Invited] Aerial Display by Use of Plane-Symmetric Real-Image Forming Optics and Its Modulation Transfer Function	山本裕紹 (宇都宮大学), 滝山和晃 (宇都宮大学), 正岡顕一郎, 陶山史朗 (宇都宮大学)	SPIE Photonics West 2024 (SPIE OPTO)	Paper 12909-18	2024.1.27～2.1
イマーシブコンテンツ体験に向けたディスプレイ技術	藤崎好英	機能性フィルム研究会	< <a href="https://kinouseifilm.com/katudou/2023-2/">https://kinouseifilm.com/katudou/2023-2/</a> >, (accessed 2024-03-11)	2024.1.31
LDM方式を適用した衛星放送システムにおける非線形補償の一検討～DPDとパイロット信号を用いた伝送性能向上～	小島政明, 小泉雄貴, 亀井 雅	電子情報通信学会技術研究報告 SAT 衛星通信	Vol.123, No.377, SAT2023-69, pp.55-59	2024.2.8～9
[依頼講演] 論文の執筆から査読まで～地上放送高度化方式の研究開発に関する論文の実例～	朝倉慎悟	映像情報メディア学会技術報告	Vol.48, No.5, BCT2024-30, p.39	2024.2.15～16
[依頼講演] ニュース要約映像作成支援システムの試作	望月貴裕	映像情報メディア学会技術報告	Vol.48, No.5, BCT2024-31, p.40	2024.2.15～16

#### 編集委員会

編集長	村山 研一
委員	相原 聡／本田 悠葵／奥田 誠／川本 潤一郎 中島 奈緒／田高 礼子／小出 大一／岸田 雅彦
事務局	木下 延博
幹事	田中 克／絹山 裕子

#### NHK技研R&D No.197 (2024年春号)

2024年5月15日発行

編集・発行	日本放送協会 放送技術研究所 ©2024 日本放送協会 〒157-8510 東京都世田谷区砧1-10-11 電話 03-3465-1111 (NHK代表) ホームページ <a href="https://www.nhk.or.jp/str/">https://www.nhk.or.jp/str/</a>
制作・印刷	株式会社 NHKビジネスクリエイト 〒150-0047 東京都渋谷区神山町5-5 NRビル1F 電話 03-3469-8230

※本誌は、「著作権法」によって著作権等の権利が保護されている著作物です。  
※本誌に掲載されている会社名・製品名は、一般に各社の商標または登録商標です。

NHK