

■ トピックス

- ・NHK受信実態調査結果 —2019年度—から
- ・NESラボの技術
- ・NESの技術の歴史 第2回 高精細表示に向けた技術開発

■ NHK R&D紹介

- ・スポーツ競技の状況を体感できる触覚インターフェース技術
- 公開されたNHKの発明考案
- NHK技研最新刊行物

■ テクノコーナー

- ・ニューラル機械翻訳システムのための日英ニュース対訳コーパスの開発

トピックス

NHK受信実態調査結果—2019年度—から

受信実態調査とは、テレビ・ラジオの放送受信におけるより良い受信環境の確保と、望ましい受信システムの確立を図るための基礎データを得ることを目的として、1949年（昭和24年）から開始し1996年までは3年ごと、そのあとは毎年実施している調査です。2019年度で42回目を迎えました。2019年度は、新4K・8K衛星放送に対応した受信システムの普及に向けた基礎データ収集を重点調査項目として実施しました。調査は2019年4月末のNHK放送受信契約世帯から抽出した11,130世帯を対象として、2019年8月～10月に実施しました。事前に調査票を郵送のうえ、調査員による面接・宅内調査を実施し、3,743世帯の有効標本を得ました。NHKは「NHK受信実態調査結果—2019年度—」として結果をまとめ、冊子を発行しています。以下に毎年実施している調査項目を中心に紹介します。

メインテレビの種類

メインテレビの種類別の所有世帯比率を示します。メインテレビが3波共用デジタルテレビに対応している世帯比率は99.2%でした（図1）。

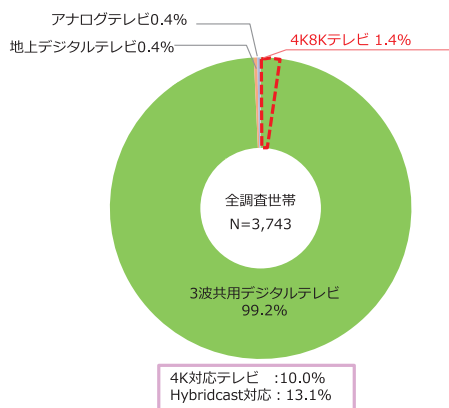


図1 メインテレビの種類

メインテレビの画面サイズとその推移

メインテレビの画面サイズの推移を示します。メインテレビの28.5%が「41型以上」でした（図2、図3）。

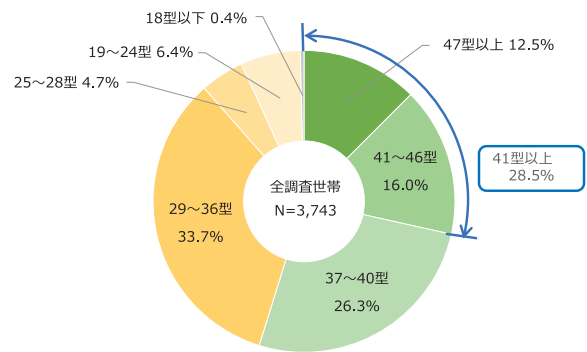


図2 メインテレビの画面サイズ

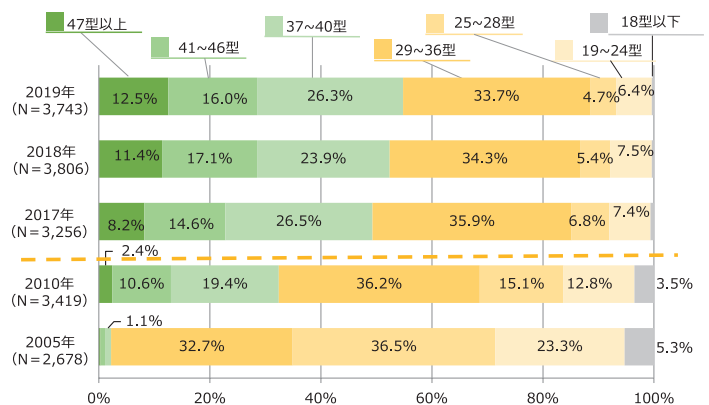


図3 メインテレビの画面サイズの推移

メインテレビの製造年

メインテレビの製造年を示します。メインテレビの製造年が10年を超えたテレビの所有世帯比率は15.4%でした(図4)。

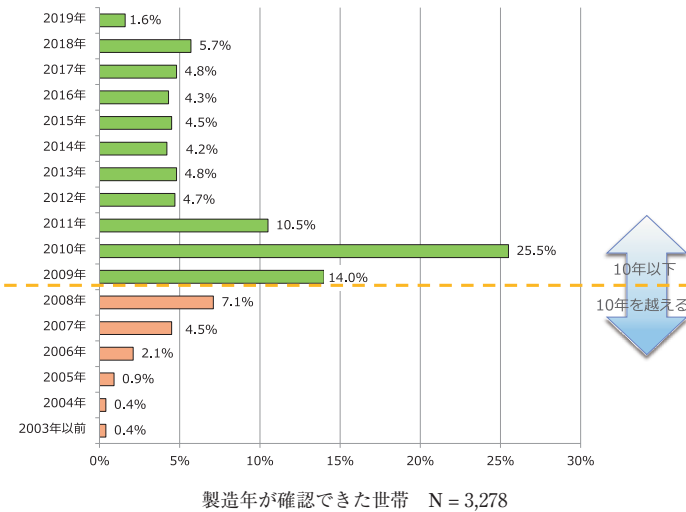


図4 メインテレビの製造年

録画機器の所有状況

録画機器を所有している世帯は73.7%でした(図5)。

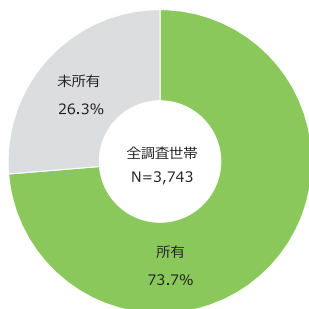


図5 録画機器の所有状況

地上放送の受信方法

地上放送の受信方法を示します。個別にアンテナを設置して地上放送を受信している世帯比率は34.0%であり、共同アンテナで地上放送を受信している世帯比率は66.0%でした(図6)。

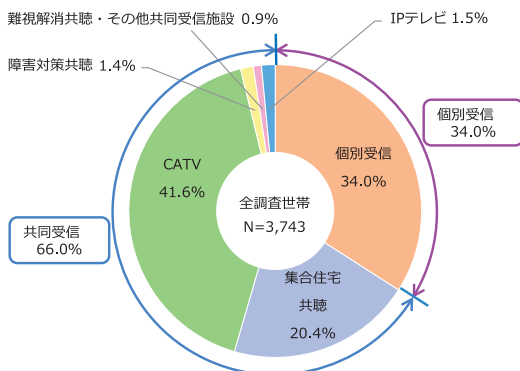


図6 地上放送の受信方法

地上放送の受信方法の推移

地上放送の受信方法別の世帯比率の推移を示します(図7)。

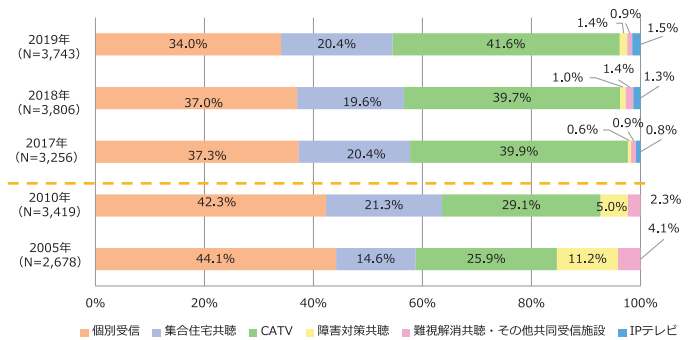


図7 地上放送の受信方法の推移

BS放送の受信率の推移

BS放送の受信率の推移を示します。「BS放送受信」の受信率は62.2%でした(図8)。

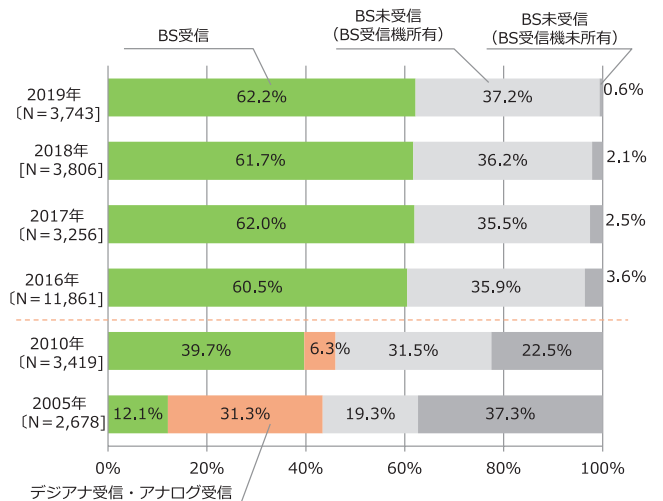


図8 BS放送の受信実態の推移

BS放送の受信方法

BS放送を受信している世帯における受信方法別の世帯比率を示します。「個別受信」の世帯比率は25.6%でした(図9)。

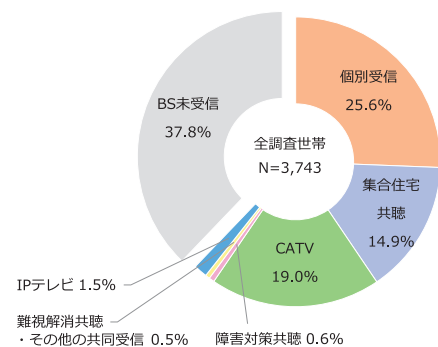


図9 衛星放送の受信方法

BS放送の受信方法の推移

BS放送の受信方法別の世帯比率の推移を示します（図10）。

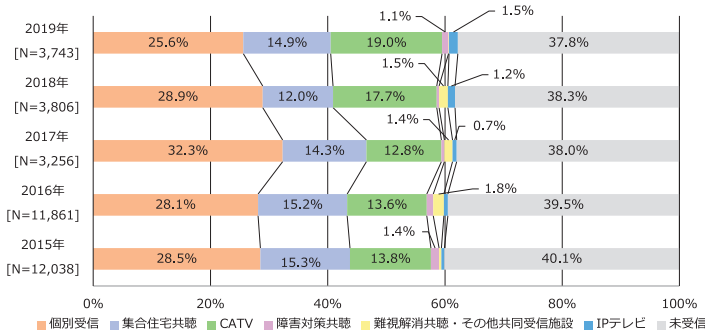


図10 BS放送の受信方法の推移

メインテレビの視聴環境

メインテレビを設置している部屋の大きさの世帯比率と、メインテレビの画面サイズと部屋の大きさの世帯比率を示します（図11、12）。

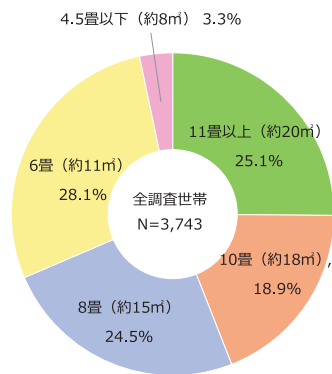


図11 メインテレビを設置している部屋の大きさ

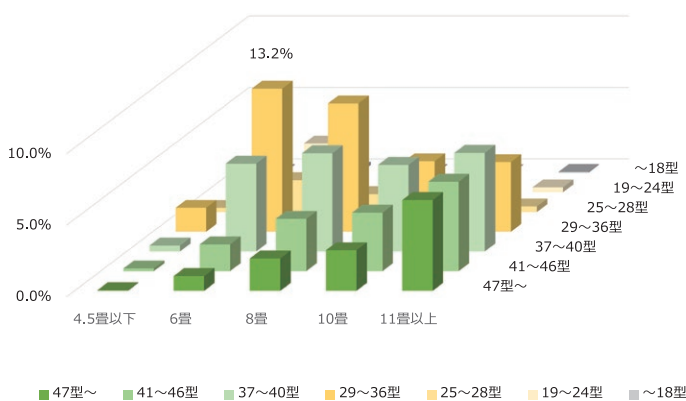


図12 画面サイズと部屋の大きさ

インターネット接続機器

インターネット回線の種類別利用状況を示します（図13）。インターネットに接続し得る受信機及び受信機以外の機器別の世帯比率を示します（図14）。

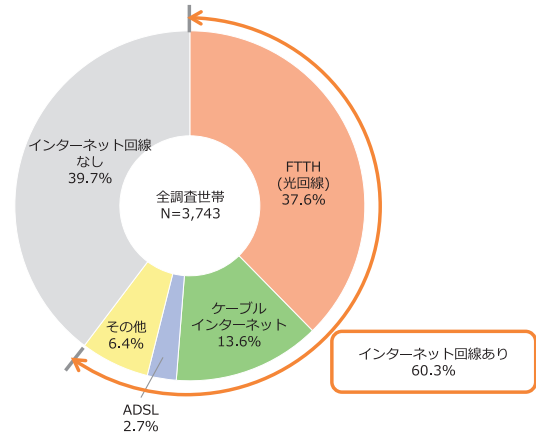
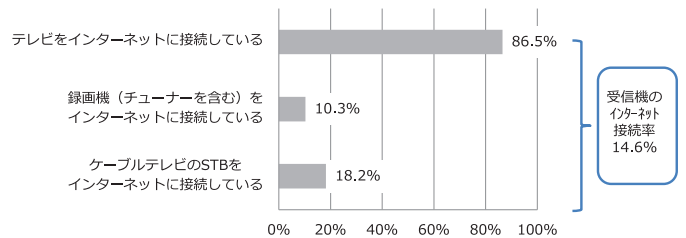


図13 インターネット回線の種類別利用状況



受信機をインターネットに接続している世帯 N = 509 (複数回答)

図14 受信機をインターネット接続している世帯

受信機のインターネット接続率の推移

受信機のインターネット接続率の推移を示します（図15）。

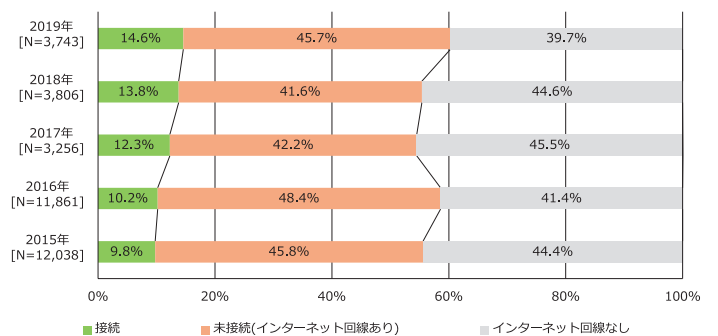


図15 受信機のインターネット接続率の推移

NHK技術局 送受信技術センター

企画部 松下 純也

(一財) NHKエンジニアリングシステム

R&T技術部 藤井 康人

NESラボの技術

——音声合成技術

2017年に和歌山県の民間ラジオ局で、大手IT企業が提供するAI（人工知能）音声合成サービスを用いたニュース・天気予報の自動放送が開始されました。ニュースのような任意文章を発話させるには、読み誤りや韻律（アクセントやイントネーションなど）の不自然さの解消がまだまだ難しいと考えられていた中でこの話題は、音声合成技術の進化を広く印象付けたとともに、放送ビジネスの新たな可能性を示すことにもなりました。

音声合成技術の研究は、一通りの発話ができるようになってからが本当の勝負だと言われます。「もっと明るい声がいい」「ゆっくり落ち着いた話し方にできないの？」など、話し方に関する要望がすぐにやってきます。「この声はイヤ」と思われてしまうとサービスそのものが成り立たなくなることもあり、声の多様性への対応は音声合成ビジネス成功の一つの鍵ということが出来ます（図1）。

そこでNESラボでは、NHKが開発した汎用的な音声合成技術をより多くの方に使っていただけるよう、ニーズに応じた声を効率的に提供するための基礎的な検討に着手しました。本稿では、当ラボでの今後の取り組みを展望する観点から、これまでのNHKでの音声合成技術の実用化事例をいくつか紹介したいと思います。



図1 話題に応じて話し方も変わります

ラジオ第二放送の株式市況と気象通報

株式市況は2010年3月から、気象通報は2016年3月から合成音声による放送が行われています。いずれの番組も典型的な文章パターン（テンプレート形式）で構成できることが特徴です。テンプレート化できるということは発話のバリエーションを見込むことができるということであり、合成すべき音声が有限の音の特徴の組み合わせで表現することができるということです。そこで、想定される言い回しを必ず含む音声データベースを予め用意しておき、その中の音声波形を適切に組み合わせることで放送品質の音声合成を実現しました。

しかし、このシステムで品質良く読むことができるのは、録音した話者の特定の言い回しに限られてしまうため、サービスの拡張性に課題がありました。例えば、株式市況で新規銘柄が追加された場合、その銘柄名は想定外の言い回しになるため上手く読むことができず、新たにアナウンサーの発話を収録して音声データベースに追加する必要があります。このような課題に対し、NESラボではNHKが開発した汎用的な音声合成技術、DNN（Deep Neural Network）音声合成技術（図2）を用いて、企業名のような文章ではない単語を品質良く読み上げることなど、自動放送システムの永続的な運用に資する要素技術の検討に着手しました。

ラジオ気象情報番組の自動放送化

2019年度中に甲府放送局と新潟放送局のラジオ第一放送で、音声合成によるラジオ気象情報のトライアル放送が実施されました。気象庁から配信される気象データから自動で読み原稿を作成し、NHKが開発したAIを活用した音声合成技術で読み上げました。アナウンサーが過去

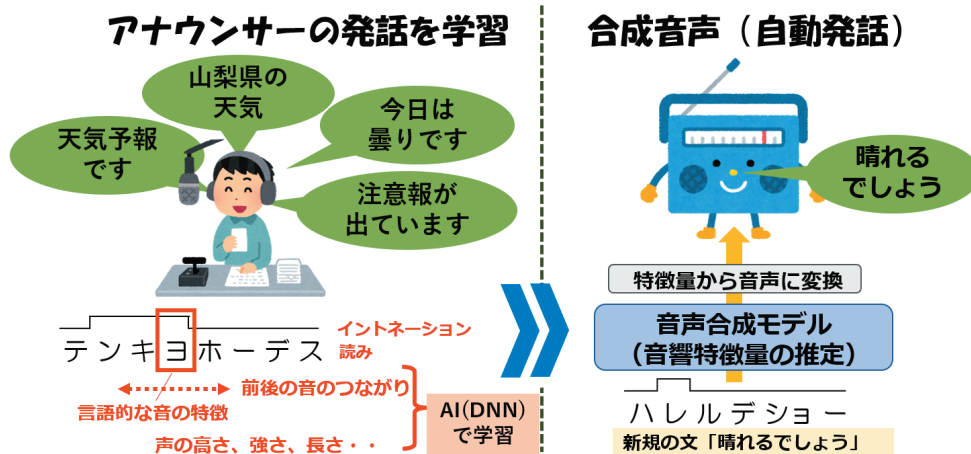


図2 DNNを用いた音声合成方式の概念図

に読んだ大量の気象情報から、文脈に応じたイントネーションや間の取り方などを学習し、アナウンサーらしい合成音声で再現しました。実際の放送では、アナウンサーは最新の気象データ（警報、注意報、予報、雨量、気温など）から、伝える情報の優先順位を考え、放送時間に収まるように原稿を組み立てています。このような作業も自動化されました。2020年度以降は複数局での本放送の実施が予定されています。

NESでは、上記システムの本放送整備に際して、自動で原稿を生成する発話生成部と音声合成部で構成されるエンジン部を担当します。NESラボでは同技術による放送サービスの充実・強化に向けて、発話の高度化や多様性、音声合成の話者の拡充などに取り組むとともに、地域局ならではの個別のニーズにも応えていきたいと考えています。

放送サービスのすそ野を広げる

—自動解説放送とロボット実況—

視覚に障害のある方のために、映像を見なくては伝わらない情報を副音声で補完する「解説放送」が行われています。しかし、その対象はドラマなど一部の収録番組にとどまっており、スポーツなど生放送番組には対応できていないことが課題でした。そこでNHKでは、五輪などの大規模なスポーツ大会でリアルタイムに配信される競技データ（いつ、誰が、何をしたなど）から解説放送的な情報を自動生成して、これを合成音声で適時発話させる、新しい考え方の解説放送サービス「自動解説放送」の研究に取り組んでいます。

このサービスにおける音声合成技術の課題としては、放送実況と同時に聴取するため、解説発話はできるだけ端的であること、実況の声と区別しやすいこととされています。一口に区別と言っても口調や声質、話速など要素はさまざまであり、かつ、実況や背景音の状況も時間的に変化していくため、その実現には多くの課題があると考えられます。NESラボではこのような課題の解決に向けて、実際にサービスを作る観点から関わっていきたいと考えています。

また、同じ技術を用いて、実況のない五輪動画に競技イベントの説明発話を付与したサービスが「ロボット実況」と命名され、放送を補完する世界初のサービスとして実用化されました。2018年2月の平昌五輪では、アイスホッケー、リュージュ、スケルトン、ボブスレー、カーリングの一部試合で、音声合成による自動生中継がインターネット

トおよびハイブリッドキャストで実施されました（図3）。図中の字幕はリアルタイム競技データから自動で生成した実況原稿です。これを音声合成で読み上げました。

このサービスにおける音声合成の課題は、スポーツ実況にふさわしい口調であること、競技イベントに対して音声合成処理が遅れないこと、さらに外国選手名を正しく読むことなどがあります。ロボット実況は今後もサービスの拡充が期待されていますが、NESラボでは、声質や口調の多様化など、よりサービスの質を高める声の提供に取り組んでいきたいと考えています。



図3 音声合成技術の応用例“ロボット実況”

今後に向けて

音声合成技術の進歩に伴い、放送サービスの中にもさまざま導入される時代になりました。放送は映像と音声によって作られた文化であり、その半分を担うことができる技術が世の中に出てきたかと思うと、一音声研究者として感慨もひとしおです。冒頭の和歌山のラジオ局の例では、自動放送の導入費用が従来の局内システムの整備に比べて1/100以下になったなどコスト面でのメリットも大きくうたわれていました。アナウンサーが時間的に拘束される定型反復業務を自動放送に置き換えることができれば、その時間を取材などよりクリエイティブな仕事にシフトすることができることから、要員の少ない地方局の視聴者サービスの向上や、働き方改革にも資することが期待されます。

アナウンサーの代替としての音声合成技術にはいずれも高い品質が求められますが、このような仕事に携わる中で、技術としての理想を追求しつつ、それを広く一般化する部分でNESラボが貢献できればと考えています。

(一財) NHKエンジニアリングシステム

先端開発研究部 上級研究員 今井 篤

NESの技術の歴史 第2回 高精細表示に向けた技術開発

前書き

前回に引き続き当財団における高精細表示関連のお話になりますが、今回はここ約10年間において当財団が開発に関わった高精細表示関連の技術を紹介します。

どのような技術開発でも似た状況でしょうが、原理的に動作すると理解していてもそれが実際に動作するのを見る喜びは格別です。新しいものを作る楽しみを一緒に味わっていただけたら幸いです。

8K小型プロジェクター

最初に紹介する技術開発は前回同様に表示装置で、e-shift素子を用いた8K小型プロジェクターです。2010年度に試作され（2011.1.13に報道発表）2012年のロンドン五輪パブリックビューイング（日本国内での会場は渋谷、秋葉原、福島）など多くの場所で使用されました。

これ以前の8Kプロジェクターは前回説明の通り4K解像度のLCOS素子を4枚組み合わせたもので、プロジェクターユニットが2台（GのLCOS素子2枚含むユニットとR、B用素子各1枚のもの）になります。原理上8K解像度があるのはG色でR、Bについては4K解像度です。プロジェクターユニットが2台あるため、大掛かりな筐体機構が必要、精密な設置調整が必要、経時変化（温度、位置など）に対応するため定期的な調整が必要、などの課題があり、数多くの展示などに対応する上で障壁になっていました。

この解決にはプロジェクターユニットを1台にする必要があります。当時利用可能な素子は4K解像度まで、通常のプロジェクター構成で1ユニットに組み込める表示素子は3枚まで、との制約がある中で考案されたのが図1に示すようにe-shift素子を用いた構成です。e-shift素子というのは、フレーム単位で光の偏光方向を変える（例えば、縦方向⇄

横方向）液晶素子と、偏光方向によって入射光の進行方向が変化する（したがって出力光の位置がずれる）複屈折素子を組み合わせたものです。複屈折素子の厚みを適切な値にして、1フレームごとにe-shift素子の液晶素子を駆動することで、図2に示すように1フレームごとに4K解像度で0.5画素だけ表示位置を移動しますので、静止画ではRGBともに等価的に8K解像度（動画では4K解像度）での表示を可能にします。プロジェクターの外観を図3に示しました。

短時間で表示位置をシフトすることによる高解像度化というのは昔のインタレースと同じではと思った方、その通りです。必要は発明の母、昔からのアイデアでも新しい技術と組み合わせると可能になるものがあるのです。1フレームごとに表示位置を移動させるので細かいパターンではチラチラとフリッカー的に見えるというのが心配になる方もいるでしょうが、GとRBで偏光方向を反対にすることでフリッカーを抑えていて（図2で、あるフレームではGがグレーの位置でRBが枠の位置、次のフレームではそれが反対になる）、これが特許（NHK、JVCケンウッド、当財団の共同出願）になっています。

この項の最後に、このプロジェクターによって可能になったものを紹介します。140インチのリアプロジェクターで、NHK放送技術研究所のロビーに2014～2016年初めまで設置されていました。内部構成が図4、外観が図5です。通常の鏡では高解像度化が難しくさらにゴースト像が出ますので、表面精度を高めた表面反射鏡を用いて

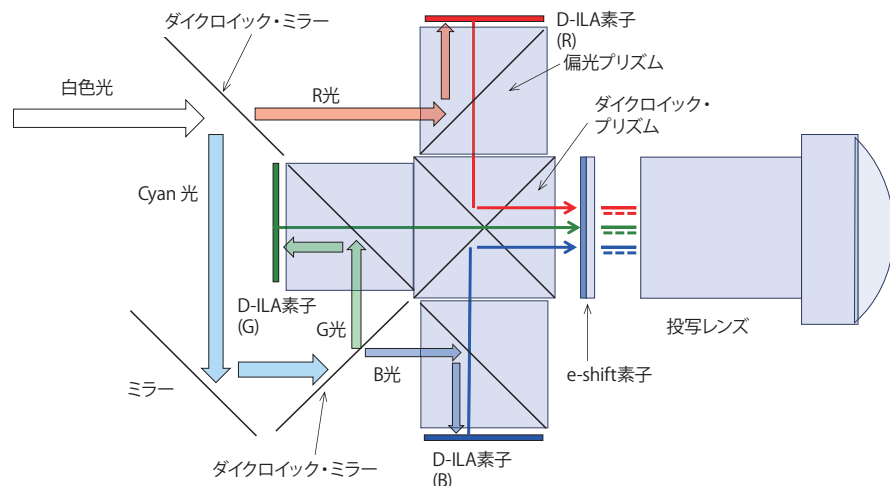


図1 8K小型プロジェクターの構成

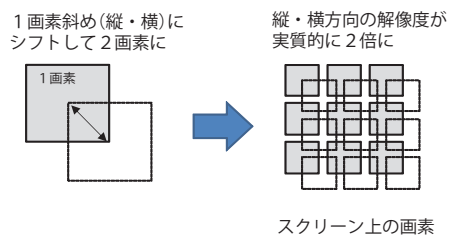


図2 e-shift素子の働き

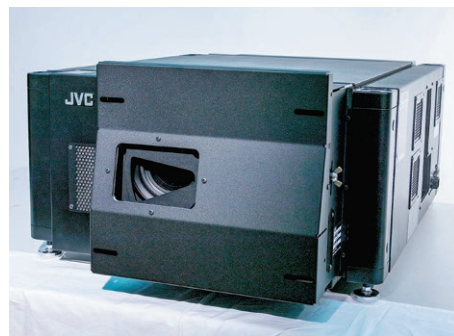


図3 8K小型プロジェクターの外観（提供NHK）

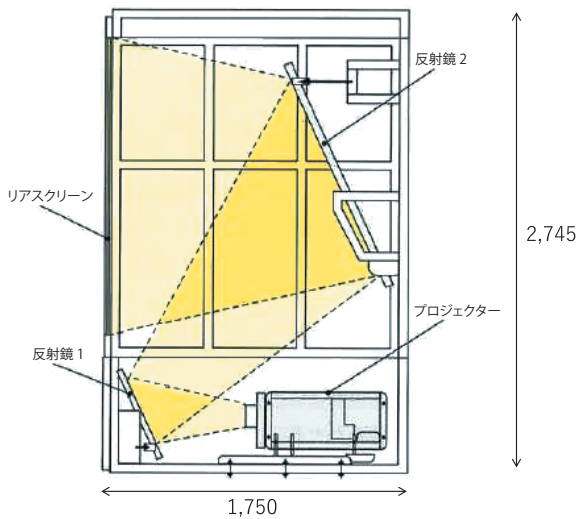


図4 8Kリアプロジェクターの構成



図5 8Kリアプロジェクターの外観

います。このような構成は、プロジェクターユニットが2台では不可能でした。

8K HDMI2.0信号変換装置

スーパーハイビジョンは2016年に試験放送が開始され今は本放送ですが、スーパーハイビジョン映像信号は従来のハイビジョン信号と比較して何が違うのでしょうか？画素数が、と言う前に色とかHDR (High Dynamic Range)と言われた方、苦労していますね。NTSCからハイビジョンになるときに規格上では3原色点が多少変化しましたが、スーパーハイビジョンでは大幅な変更になりますし、HDRというものとはどちらかというところ最近発達したものです。一般に映像信号の形式が変わるときには何らかの変換装置が

必要になります。前置きが長くなりましたが、次に紹介する当財団の技術開発は8K HDMI2.0信号変換装置です。

放送される映像信号はITU-R規格で規定されていますが、スーパーハイビジョンの基準となる3原色点の規格のBT.2020が成立したのは2012年、HDRの規格のBT.2100は2016年です。スーパーハイビジョン放送はこれらの規格が規定する映像信号で行われるのに対し、少なくとも2016年以前に開発された機器はこれらの規格には対応していませんので、スーパーハイビジョン放送に対応するためには信号変換装置が必要になるのです。構成としては、図6に示すように、3原色変換をするリニアマトリクスと光電気変換の補正を行うLUT (Look-Up-Table) の組み合わせで対応できます (3入力3出力となる他の方法でも可能です)。8K HDMI2.0信号変換装置の場合は、スーパーハイビジョン放送のデコーダー出力がHDMI信号、従来の主なスーパーハイビジョン用機器がHD-SDI信号なので、HDMI信号からSDI信号への変換も含まれます。

ここで光電変換に関して補足します。HDRというものが出てきたため従来の方式はSDR (Standard Dynamic Range) と呼ばれます。スーパーハイビジョン放送ではSDR、HDR両方が可能なため図6ではどちらにも対応できる構成になっています。なおITU-R BT.2100規格ではHDRとしてPQ (Perceptual Quantization) 方式とHLG (Hybrid Log-Gamma) 方式が規定されていますが、日本の放送で用いられているのはHLG方式です。

図6の構成は目新しいものではなくリニアマトリクスやLUTは規格から求められます。8K HDMI2.0信号変換装置は映像機器メーカーの計測技術研究所が製造しました。それでは、当財団はこの開発で何を担当したのだろう、と不思議に思われる方もいるかと思いますが、今回は全体構成とリニアマトリクスやLUTの係数および動作テスト時のテスト信号などのチェックリスト制作を担当しました。映像機器メーカーといえども得手不得手があります。このようにお互いができることを持ち寄って開発するということが必要になる場合はあるのです。

頻度は徐々に低くなっていますが規格に対応していない機器を使用する機会はあるので、8K HDMI2.0信号変換装置はまだ元気に活躍しているようです。

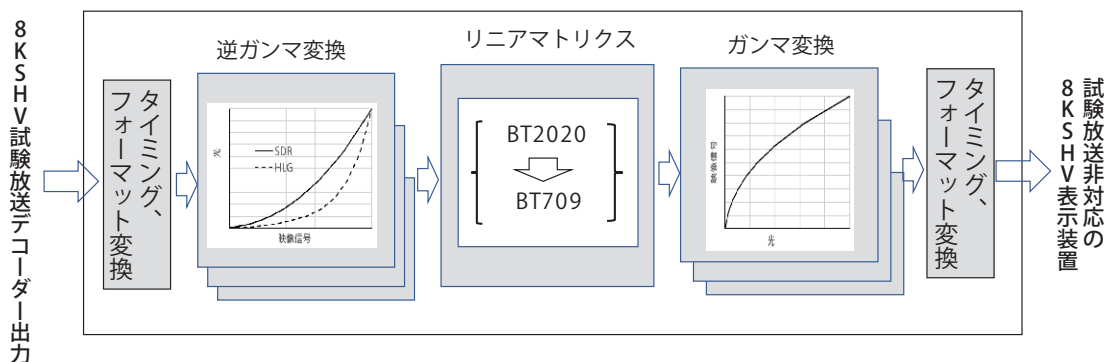


図6 8K HDMI2.0信号変換装置の構成

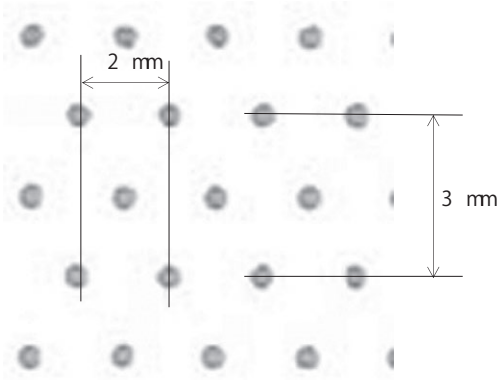


図7 映画館のスクリーンの拡大(例)

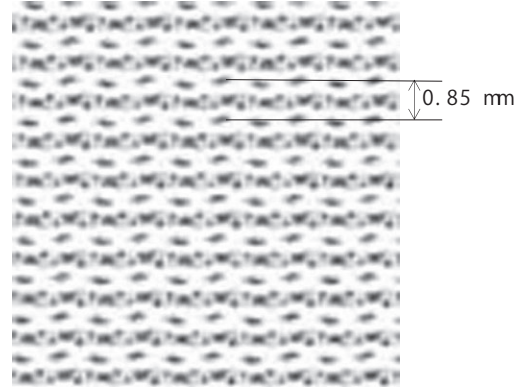


図8 開発した音響透過スクリーンの拡大

音響透過スクリーン

最後に紹介する研究開発は、8Kプロジェクター用の反射型の音響透過スクリーンで、8K用ダビングスタジオなどで使用されています。まず音響透過スクリーンについて簡単に説明しましょう。映画館のスクリーンを近くで見ると多数の微小な点が見えるはずですが、これはスクリーン裏側に設置したスピーカーの音を通すための穴で、スクリーン生地にレーザーを照射して開けているそうです。この音を通す音響透過の機能により視聴者は画面方向からの音を聞くことができます。これがない場合は、スピーカーを画面の上下において合成するしかない（ファントムスピーカーと呼ばれます）ので、音の方向は正確には再現できません。

8K用ダビングスタジオでは8Kの22.2マルチチャンネル音響での番組制作のため、規格（BS.2051）でのスピーカー配置と8K画像表示が求められ、8Kプロジェクターと音響透過スクリーンが必要となりました。映画館で使用される音響透過スクリーンをそのまま用いればいいと思った方、世の中はそんなに甘くありません。図7に映画館で使われるスクリーンの拡大を示すように、2~3mmピッチで穴が開いています。これに対し8K用ダビングスタジオの画面サイズは220インチ、約3m離れた位置で画面を見ます。この条件での8K画像の1画素は0.64mm（前項の8K小型プロジェクターを用いる場合は4K画像での位置ずらしなので倍の約1.3mm）です。図7の穴のパターンは1画素の数倍と粗く目立ちやすく画像に対しては妨害となりますので、別な構造の音響透過スクリーンが必要となります。

画像への妨害を少なくするためには穴のピッチを画素サイズ程度に小さくすればいいのですが、画素サイズに近い大きさの場合は画素構造と穴のパターンでの干渉のため画面上にモワレと呼ばれる縞模様などが生じます。モワレについて補足します。干渉ですので画素構造のプロジェクターと音響透過スクリーンを使用する限りモワレは生じますが、モワレが目立つかどうかは画素サイズと穴のピッチの関係、穴の構造などによって変化します。このためスクリーンサンプルなどでの事前テストが必要となります。

今回開発したスクリーンの拡大を図8に示しました。これは細かい糸での織物に光を反射させる材料をコーティングしたもので穴のピッチは0.85mm、イーストンが製造しました。当財団がこの開発の何処にどう関わったのだろう、という試作に伴うモワレなどの評価です。お互いの得手不得手をカバーすることでの共同開発です。

開発したスクリーンは織物なので織機の大きさで織物の大きさには制限が生じますが、つなぎ目を目立たなくする手法を開発しましたので（NHK、イーストン、当財団が共同で特許出願中）より大きなスクリーンサイズにも対応できるようになっています。これからも8Kスーパーハイビジョンの音作りを支える製品として使用されていくものと期待します。

終わりに

最近10年間ほどでの高精細表示関係での技術開発を紹介しました。当財団はメーカーではありませんので機器などの開発は他の会社とお互いができることを組み合わせてのものになります。お互いの得手不得手をカバーしあう組み合わせは、やってみると結構ありますし、少し違う会社同士の交流が刺激になり次のステップへの弾みにもなります。

このような開発は、相手の会社だけでは難しいことに当財団が関わることでニーズに合わせたものを作るものです。このことから当財団自身は、相手の会社と現場のニーズをつなぐ一種のインターフェースといえるでしょう。

今回のウイルス問題で映像現場などでのニーズは変わっていくでしょう。当財団がその変化に適合したインターフェースであり続けるよう努力しますので、今後ともご期待ください。

金澤 勝*

* 執筆時、(一財)NHKエンジニアリングシステム 先端開発研究部 研究主幹

ニューラル機械翻訳システムのための日英ニュース対訳コーパスの開発

はじめに

当財団は情報通信研究機構の委託研究「多言語音声翻訳の高度化のためのディープラーニング技術の開発」の下で行われているニュースの日英機械翻訳システムの研究プロジェクト（以下、ニュース翻訳プロジェクトと略記）に参加し、高性能なシステムを目指した研究開発を進めています。

ニューラルネットワークにより画像認識の性能が飛躍的に向上したことをきっかけに、音声認識、音声合成、文書要約、文書生成などたくさんのアプリケーションにニューラルネットワークが適用され、性能の大幅な改善が得られています。機械翻訳も同様です。ニューラルネットワークによりさまざまな言語対、分野で翻訳性能が飛躍的に向上することが確認され、さらなる性能向上、実用に向けた研究開発が盛んになっています。

ニューラルネットワークの利用の鍵の一つが、大量かつ高品質な学習データの入手です。どのような応用であってもニューラルネットワークを応用したシステムの性能は学習データによって決まると言っても過言ではありません。つまり高性能なニュースの日英翻訳システムの実現には、何十万、何百万という規模のニュース文の日英対訳「日英ニュース対訳コーパス」が必須となります。

残念ながら現時点で一般に利用できる大規模な日英ニュース対訳コーパスはなく、高品質なニュースの翻訳システムは実現できていません。そこでニュース翻訳プロジェクトでは参加機関の一つ、株式会社時事通信社の

日本語と英語のニューステキストを元に高品質な大規模日英ニュース対訳コーパスを開発することにしました。以下では当財団が主に担当しているコーパス開発の概要と予備実験の結果を紹介します。

時事通信社のニュース記事を使ったコーパス開発

時事通信社は日本語のニュースを内外の報道機関や一般読者に提供しています。また日本語ニュースの一部を英訳して同様に提供しています。この日々の業務で作成された日本語と英語の記事を提供してもらいました（表1）。表1より日本語記事の3.7%程度が英訳されていることが分かります。私たちは、日本語と英語の記事のペア「日英対応記事」と日本語のみの記事「日本語単独記事」の2つのデータからコーパスを作ることにしました。

表1 コーパス作成に利用した記事

日本語記事	1,561,143
英語のある日本語記事	57,154

日英対応記事を利用したコーパス作成

日英対応記事には記事同士の対応があるため、文の対応を取れば対訳コーパスを作ることができます。ここで、平均的な日英記事を観察して特徴を見てみましょう（図1）。

図1の左は日本語記事、右はその英語記事です。わかりやすいように英語記事を日本語に直訳しています。また対応する文同士を線で結んでいます。冒頭はタイトル

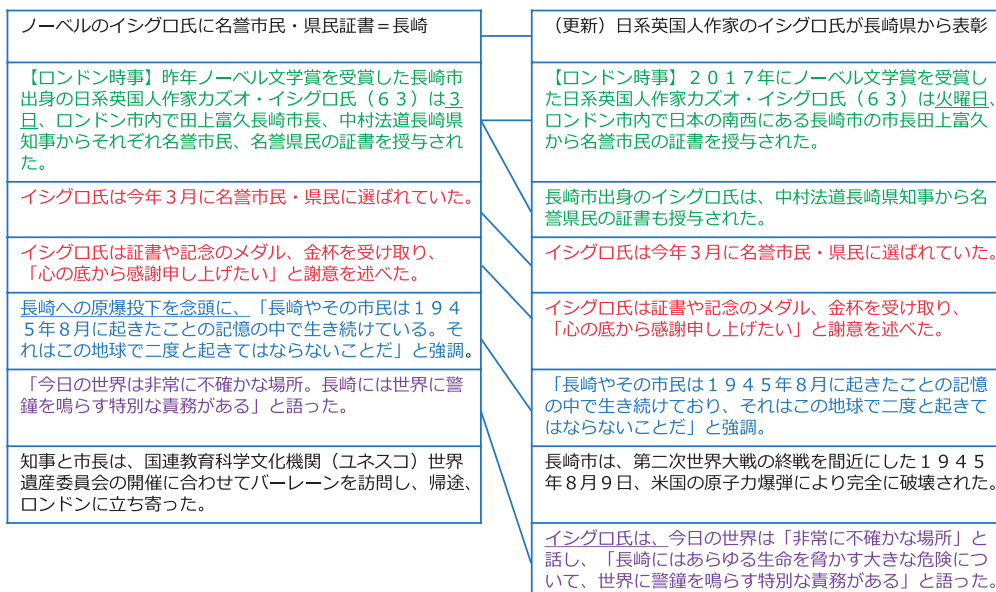


図1 日本語記事と英語記事（日本語に翻訳）

で、その下に日本語記事の本文が6文、英語記事の本文が7文あります。以下、本文を詳細に見ていきます。

日本語の第1文は英語の第1文と第2文に翻訳されています。内容はほぼ同じですが、日本語の「3日」という日付が英語では「火曜日」という曜日で翻訳されています。日本語の第2文と第3文は英語の第4文と第5文にそのまま翻訳されています。日本語の第4文は英語の第5文に翻訳されていますが、冒頭の「長崎への原爆投下を念頭に」というフレーズが英語では省略されています。その代わりに英語の第6文に1945年の長崎への原爆投下の事実が説明されています。原爆投下の説明は日本語記事にはなく、英語読者に配慮した追加だと考えられます。日本語の第5文は英語の第7文にそのまま翻訳されています。ただし日本語で省略されている主語「イングロ氏は」が英語に現れています。日本語では主語の省略が可能ですが、英語では必須なためこのような違いを生じています。最後の日本語の第6文は翻訳されていません。

以上の観察から、一文が複数の英文で翻訳されることがある、日英片側にしかない文が存在する、日英で対応する文であっても片側にのみ出現する表現がある、といった特徴を把握できます。また、本例にはありませんが、対応する文の順序が日英で逆転する場合があります。つまり英語記事は日本語記事の忠実な翻訳ではありません。

上記の日英ニュースのような2言語文書の文対応を自動推定する文アラインメントと呼ばれる手法があります。文アラインメントは対訳辞書などを使って2言語ニュースの単語対応を求め、単語対応を元に文対応を推定します。図1のニュース記事の場合、文アラインメントが成功すると文間の線で示された正しい文対応が得られます。文アラインメントを使うと日英対応記事からコストをかけずに対訳コーパスを作成できるため、私たちもこの手法でコーパスを作成しています。しかし、次のような問題を生じます。

・部分的な対応

文アラインメントは、文対応の候補と日英の文の類似度を出力します。通常、あらかじめ設定された類似度の閾値以下の対応を不採用とします。つまり捨てられる対応が出てきます。また図1のように、そもそも対応しない文もあります。すなわち、記事の一部からしか文対応を抽出できないという部分性の問題を文アラインメントは抱えています。

・不均衡な情報

図1に示したように、対応付けられた日英の文には片側にしかない情報が含まれることがあります。また、日付と曜日のように、対訳として不正確な情報も含まれます。これらの不均衡な情報を含む対訳を学習に使うと誤訳、訳抜け（原文の表現が翻訳に現れない）、過剰訳（原文にない表現が翻訳に現れる）の原因になります。

不均衡な情報を減らすには、文対応の採用を決める類似度の閾値を大きくすると効果がありますが、得られる

文対応は減ってしまいます。適切な閾値の設定は難しい課題です。

日本語単独記事を利用したコーパス作成

文アラインメントに加えて、日本語単独記事の人手英訳によって対訳コーパスを作成しています。コストはかかりますが、文アラインメントの持つ問題は発生せず、高品質なコーパスを構築できます。しかし、人手の翻訳にはいくつかの課題があるため、私たちは以下の方法で翻訳を進めました。

・記事選択

表1の日本語記事全てを翻訳することはコストの面で困難です。また実際、3.7%程度の日本語記事しか翻訳されていないことから、以下の方針で記事を選択しました。日本語記事の中には、経済指標、株、スポーツの結果など数字の表だけの記事があり、これらを翻訳対象から除外しました。

英訳されている日本語記事の80%が5文から15文の範囲にあることから、文数がこの範囲に入る日本語記事を選択しました。

日本語記事に付いているキーワードリストを使って英訳されにくい記事を検出し、除外しました。

・翻訳の方針

翻訳システムの学習に使うには1対1の文対応が最も取り扱いが容易です。しかし、長い日本語文を自然な英語1文に翻訳するのは困難です。そこで日本語の1文を英語の1文、あるいは長文の場合は複数文で翻訳しました。この時、日英の情報均等なよう内容の省略、追加は行わないようにしました。さらに、学習に使いやすいよう日英の文の順序は変更しないようにしました。

図1で説明した日本語に欠けている主語を英語に翻訳するには冒頭から翻訳対象までの文、すなわち文脈を見なくてはなりません。そこで文脈を使って翻訳するようにしました。現時点では文脈を機械翻訳システムで扱う手法は確立しておらず、研究の焦点の一つになっています。私たちのコーパスはそのような研究に使えることも目指しています。

時事通信社の英語ニュースでは特別な用語、例えばpercentの代わりにpctという略語が使われます。時事通信社では訓練を受けたニュース専門の翻訳者が英語のニュースを作成しますが、私たちは一般の翻訳者に作業を依頼します。そこで翻訳者に用語のルールを守ってもらうため、時事通信社のニュースのスタイルガイドを事前に作成して提供しました。

日英ニュース対訳コーパスの特徴

文アラインメントと翻訳者が作成した対訳コーパスの違いを調査するため翻訳実験を行いました。実験に使った2つのコーパスの諸元を表2に示します。Alignは文アラ

インメントで作成した対訳コーパス、Manualは翻訳者が英訳して作成した対訳コーパスです。

表2 実験に使ったコーパスの諸元

	Align	Manual
抽出時期	2011-2018	2016-2018
文数	239,781	199,170
日本語平均長	54.5	48.9
英語平均長	26	25.2
文長相関係数	0.599	0.896

Alignコーパスは1対1の文対応で、類似度が0.3以上の対訳を抽出して作りました。抽出前の1対1の文は58万文でしたが、類似度による抽出で約24万文に減っています。表1の日英対応記事57,154本の平均文長は「日本語47.7文字、英語25.0単語」でした。これらと比較すると、Alignコーパスの日本語文は通常翻訳される文より少し長めで、これも類似度による抽出の影響だと考えられます。

日本語と英語の内容の一致を表す指標として、日英の文長の相関係数を計算しました。日英の文長が比例する場合は日英の内容がよく対応していると考えられ、この場合、相関係数は1に近くなります。表2のManualコーパスの相関係数はAlignコーパスより1に近く、Manualコーパスは日英の情報がより一致していることが推察できます。

翻訳実験

表2のコーパスの効果を観察するためTransformerを使った翻訳実験を行いました。Transformerはニューラルネットワークを使った最近の翻訳システムで、現在一番性能が良いと言われています。

ここではAlignとManualコーパスで学習した2つの翻訳システムの性能を比較しました。特に学習データの量の

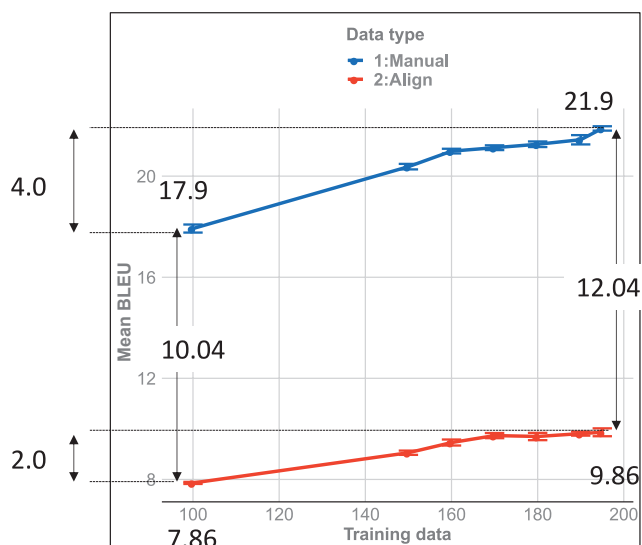


図2 学習データの量と翻訳性能の変化

影響に着目するため、学習データを10万文（100K）から19.5万文（195K）まで変化させて性能を評価しました。結果を図2に示します。

縦軸はBLEUと呼ばれる性能指標で、値が大きいほど性能が良いことを示します。横軸は学習データの文数です。Align、Manualコーパスとも学習データが増えると翻訳性能が向上しています。両者を比較すると、学習データが10万文（100K）の段階でManualとAlignのBLEUスコアの差は10ポイント程度ですが19.5万文（195K）の段階では12ポイントに広がっています。学習データが19.5万文の時の翻訳例を表3に示します。

表3 2つのシステムの翻訳例

入力	中銀は国際取引を支える上で十分な外貨準備高を保有していると話している。
Align	"The central bank has a relatively high foreign exchange reserves, the central bank said."
Manual	The central bank said that it possesses a sufficient amount of foreign currency reserve in terms of supporting international trading.

Manualの翻訳はほぼ完全に日本語の内容を反映しています。一方、Alignの翻訳は引用符の位置が誤っており、「国際取引を支える上で」の翻訳が抜けています。また、「central bank」が2回出現して冗長な印象を受けます。以上、今回の実験ではManualコーパスの方が有効だという結論が得られました。コストはかかりますが、人手の翻訳は、忠実な翻訳システムを実現するには効果が高いと言えます。

終わりに

当財団が参加しているニュース翻訳プロジェクトの日英ニュース対訳コーパスの研究開発を紹介しました。当財団では現在もコーパスの開発を進めており、研究期間の終了する2021年3月まで、できるだけ規模を拡大させていきます。また新たなコーパス構築法も試しています。さらにニュース翻訳プロジェクトでは、Align、Manualといった性質の異なるコーパスの適切な混合法、文脈を使った翻訳システムの研究も進めています。引き続き高性能な日英ニュース機械翻訳システムの実現を目指した研究開発を進めます。

謝辞

本研究は国立研究開発法人情報通信研究機構の委託研究の一環で行いました。

(一財) NHKエンジニアリングシステム

システム技術部 上級研究員 田中 英輝

スポーツ競技の状況を体感できる触覚インターフェース技術

NHK放送技術研究所では、誰もがスポーツ番組を楽しめるユニバーサルサービスの実現を目指し、競技の状況を体感できる触覚インターフェース技術の研究開発を進めています。視覚と聴覚に加えて触覚情報も提示することで、実況だけでは伝えることが難しい選手やボールの動きなどの競技の状況や、スポーツの迫力を視覚に障害がある方にも伝えられるようになります。

今回、バレーボール競技を対象に、触覚で競技の状況を伝える技術を開発しました。

システム構成

開発したシステムでは、まず、映像解析で得られたボールの位置情報のデータから、ボールがあるコートの情報を抽出し、ボールに加わる力の向きや大きさを推定します。そして、これらの情報に基づいてさまざまな触覚インターフェースを効率的に制御します（図1）。

触覚情報を提示するインターフェースとして、リストバンド型とボール型を開発しました（図2）。リストバンド型はさまざまな振動を提示できる4つの振動子を備えており、装着するとこれらの振動子が手首の上下左右に接

触します。これを両手首に装着し、例えば、右コートの選手が左コートに向かってアタックを打ったときには、右手首の左側の振動子がアタックを表すパターンで振動し、両手を塞ぐことなく競技の状況を伝えることができます。

またボール型は、アタックやレシーブの際にボールに加わる衝撃の強さを振動の大きさとともに、握った部分の側面を内蔵のサーボモーターで物理的にへこませる（凹刺激）ことで衝撃の方向も表現できます。技研公開2019では、このボール型インターフェースを用いたシステムを展示し、多くの方に体験して頂きました。

今後に向けて

今後は、競技の状況がユーザーにどの程度伝わるかの評価を進めます。また会場の歓声などを振動で表現することで、聴覚に障害がある方にもよりスポーツ番組を楽しんで頂ける技術の研究開発を進めます。

NHK放送技術研究所

スマートプロダクション研究部 東 真希子

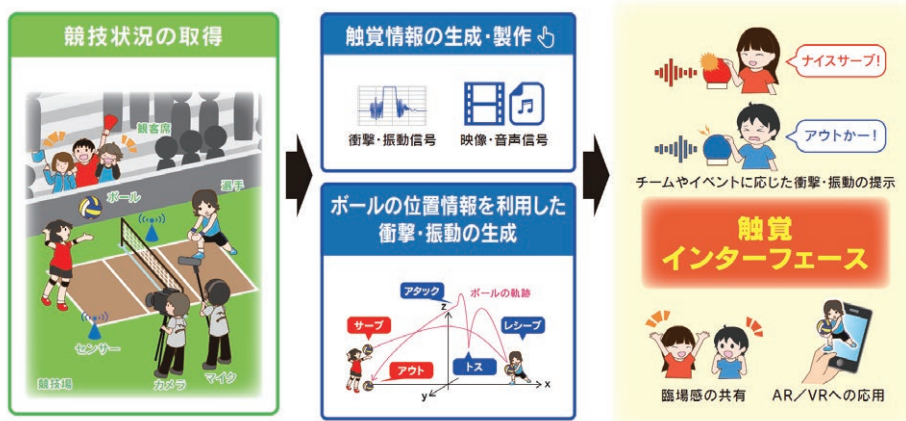


図1 スポーツ競技の状況を触覚で提示するシステムのイメージ



図2 開発した触覚インターフェース（左：リストバンド型、右：ボール型）

公開されたNHKの主な発明考案

(2020年5月1日～2020年6月30日)

発明考案の名称	技術概要
MTF測定装置およびそのプログラム 特開2020-71144	従来のエッジ法よりも計算コストを抑え、直線でないエッジからでもMTFを測定することが可能なMTF測定装置を提供するMTF測定装置およびそのプログラム
MTF測定装置およびそのプログラム 特開2020-71145	MTFを測定する対象となる領域を、エッジの傾きに応じて任意の位置や傾きで指定することが可能なMTF測定装置を提供するMTF測定装置およびそのプログラム
画像符号化装置、画像復号装置、及びプログラム 特開2020-72458	符号化効率を改善する画像符号化装置、画像復号装置、及びプログラム
触覚情報提示システム 特開2020-77094	ユーザの体の自由度を損なうことなく、映像コンテンツの内容を直観的に、触覚情報によりユーザに提示する触覚情報提示システム
映像補間装置及びプログラム 特開2020-77943	照明状態が時分割的に変化する環境下で撮影した映像を用いて、動いている物体の形状を推定するために必要な画像を生成する映像補間装置及びプログラム
信号処理回路及び撮像素子 特開2020-77992	特殊な光学系を必要とせず、画像信号の生成時にノイズを増大させることがない符号化撮像素子が可能な信号処理回路及び撮像素子を提供する信号処理回路及び撮像素子
色域表示制御装置、色域表示情報生成装置、及びこれらのプログラム 特開2020-86149	明度・彩度・色相のカラーアピランス空間情報が持つ元の色表現をそのまま維持して2次元化した色域表示情報を生成し表示制御する色域表示制御装置、当該色域表示情報を生成する色域表示情報生成装置、及びこれらのプログラム
ボット判定スコア算出装置及びそのプログラム 特開2020-86623	人間に対するプロセス及び学習データを必要とせず、ボットを正確に判定するボット判定スコア算出装置及びそのプログラム
復号装置、ホログラム再生装置、及び復号方法 特開2020-87487	多値情報を有する2次元信号から情報の復号を行う際に、データ誤りを低減することのできる復号装置、ホログラム再生装置、及び復号方法
錯視立体画像生成装置及びそのプログラム 特開2020-88496	ホロマスク錯視が発生する立体画像を生成する錯視立体画像生成装置及びそのプログラム
自動タイムキープ装置及びプログラム 特開2020-88751	スタジオ制作におけるタイムキーパーの負担を軽減する自動タイムキープ装置及びプログラム
テキスト分類装置、学習装置、およびプログラム 特開2020-91549	テキストを分類するためのモデルの学習を短時間で行える学習装置、学習済みのモデルを用いてテキストを分類することのできるテキスト分類装置、学習装置、およびプログラム
降雨減衰予測装置及びプログラム 特開2020-92367	降雨減衰を簡易かつ確実に予測する降雨減衰予測装置及びプログラム
塗布型金属酸化化物膜の製造方法、それを用いて製造された塗布型金属酸化化物膜および電子デバイス 特開2020-96134	塗布製法を用いて金属酸化化物膜を作製する際に、より簡易な製造プロセスにより形状にばらつきのない、高特性な金属酸化化物膜を形成し得る、塗布型金属酸化化物膜の製造方法、塗布型金属酸化化物膜および電子デバイス
映像合成装置及びプログラム 特開2020-96267	テレビジョンカメラの実写映像における不要被写体をコンピュータグラフィックスにより精度よく隠蔽して合成映像を生成する映像合成装置及びプログラム
パーソナルデータ利用装置およびプログラム 特開2020-98400	パーソナルデータを外部に流出させることなく、パーソナルデータを活用した各種サービス等を実現することのできる、パーソナルデータ利用装置およびプログラム
テキスト情報判定装置及びそのプログラム 特開2020-98454	話題性が高いテキスト情報を正確に判定できるテキスト情報判定装置及びそのプログラム

NHK技研最新刊行物

『NHK技研だより』

(2020年 7月号)

Top News

深さ方向で光を3原色に分離する有機撮像デバイスを開発

News

「国際規格の音響特性を実現する小型・大出力のモニタースピーカーを開発」

R&D

「耐量子性のある電子署名技術
～量子コンピューター時代の安全・安心な暗号技術」

連載 技研開所90周年記念企画 研究の歴史
(第3回／全6回)

「デジタル放送の研究」



『NHK技研だより』

(2020年 8月号)

Top News

VR・ARを活用した空間共有コンテンツ視聴システム～遠隔地の人と一緒にコンテンツを楽しむ～

News

「NAB Show Expressで研究成果を紹介」
「第31回電波功績賞を受賞」

「情報通信技術賞（TTC会長表彰）を受賞」

R&D

「有機エレクトロルミネッセンス用材料の開発～フレキシブルディスプレイの長寿命化と低消費電力化に向けて」

連載 技研開所90周年記念企画 研究の歴史
(第4回／全6回)

「8Kスーパーハイビジョンの研究」



『NHK技研R&D』182号

(2020年 夏号)

番組制作支援のための 画像・音声処理技術 特集号

巻頭言

「番組制作支援のための画像・音声処理技術特集号に寄せて」

解説

「映像自動要約技術の最新動向」

「生放送番組における自動字幕制作の最新動向」

報告

「白黒映像の自動カラー化システムの開発」

「テレビ映像を対象とした顔認識技術」

「音声認識を用いた書き起こし支援システムの開発」

研究所の動き

「地上放送高度化に向けた高効率映像符号化技術 次世代映像符号化方式VVC」

「貴重な白黒フィルム映像がカラーでよみがえる 白黒映像の自動カラー化技術」

「高画質で低遅延な4K8K映像を伝送 ミリ波スーパーハイビジョンワイヤレスカメラ」

論文紹介／発明と考案／研究会・年次大会等
発表一覧



VIEW (NHK エンジニアリングシステム友の会会誌)

Vol.39 No.5 (通巻228号)

発行日●2020年9月25日

編集・発行●一般財団法人 **NHK** エンジニアリングシステム

〒157-8540 東京都世田谷区砧 1-10-11 TEL: 03-5494-2400 (代) FAX: 03-5494-2152

制作・印刷●三美印刷株式会社

*掲載記事の無断転載を禁じます。

ITE

4K/8Kテレビシステム評価用標準動画像 Aシリーズ 頒布のご案内

一般社団法人映像情報メディア学会（ITE）は一般社団法人電波産業会（ARIB）とともに、4K/8Kテレビ放送技術の開発に必要不可欠である「超高精細・広色域標準動画像Aシリーズ」の頒布を開始いたしました。

【主な特徴】

- ・ITU-R 超高精細度テレビジョンのスタジオ規格ITU-R勧告BT.2020（Rec.2020）に準拠した動画像
- ・3300万画素CMOS 3板カメラを用いて制作した8K非圧縮映像
- ・撮影した4320/59.94Pのシーケンスからクロッピングした2160/59.94Pの4K素材もセットで提供
- ・UHDTVマルチフォーマットカラーバー（ARIB STD-B66 1.0版準拠）も提供
- ・シーケンスは、「舞妓」「着物姿の女性」「十二単の女性」画像を含む全11シーケンスで構成



仕様	Aシリーズ（8K素材）	Aシリーズ（4K素材）
画像フォーマット	7680×4320画素, 12bit, RGB 4:4:4, 59.94Hz(59.94p)	3840×2160画素, 12bit, RGB 4:4:4, 59.94Hz(59.94p)
シーケンス数	11	10
シーケンス時間		15秒
データ形式		DPX

一般社団法人 映像情報メディア学会

〒105-0011 東京都港区芝公園3-5-8 tel:03-3432-4677 fax:03-3432-4675

<https://www.ite.or.jp/content/chart/>



新4K8K衛星放送の普及を 万全の体制で支えます

BSAT (株) 放送衛星システム
BROADCASTING SATELLITE SYSTEM CORPORATION

〒151-0063 東京都渋谷区富ヶ谷1丁目16-4 パークサイド山本館
PARKSIDE-YAMAMOTOKAN, 1-16-4, TOMIGAYA, SHIBUYA-KU
TOKYO 151-0063, JAPAN TEL:03-5453-6521(代)

2019年4月1日新会社始動
～総合技術会社としてさらに進化～



NHKテクノロジーズ

最先端の放送技術 × 確かな情報システム技術

〒150-0047 東京都渋谷区神山町 4-14 第三共同ビル

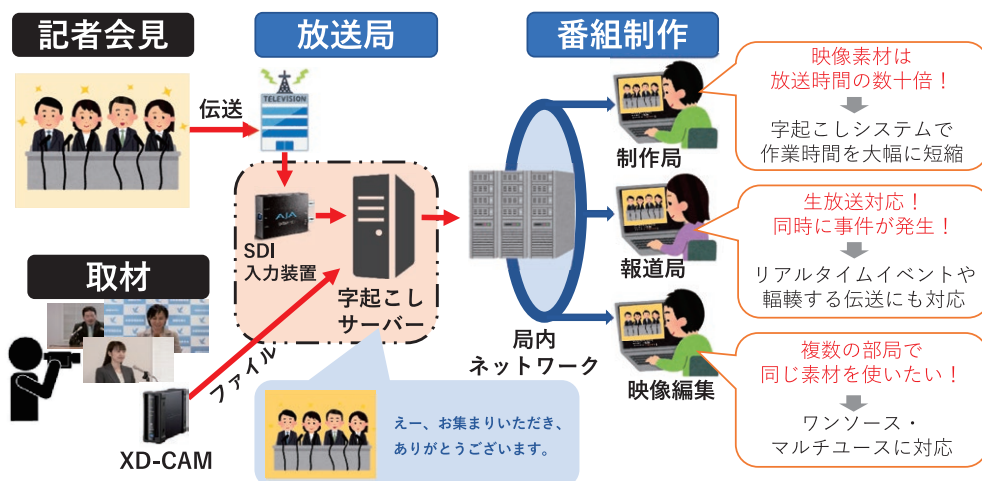
TEL:03-3481-7820 FAX:03-3481-7623 <https://www.nhk-tech.co.jp>



放送業界の働き方を変える

https://www.nes.or.jp/nes_lab/01.html

字起こしシステム



NES

一般財団法人

NHKエンジニアリングシステム

広く社会に、放送技術の可能性を届けたい

〒157-8540 東京都世田谷区砧 1-10-11
TEL: 03-5494-2400 FAX: 03-5494-2152