

スポーツ番組の自動実況・解説放送

熊野正(NHK 放送技術研究所)

1. はじめに

近年多くのスポーツ競技でリアルタイムに作成・提供されている、試合のスコア・スタッツ・イベント情報等のライブデータを利用して、試合の状況を聴いて理解できるような一連の音声発話を自動生成する研究を進めている(熊野, 2017). 人間の実況者のいない、会場音声(および映像)とあわせて聴取すれば、実況として機能するものであり、これを「自動実況」と呼んでいる. 2018年2月には、この枠組を用いて実況つきスポーツ生中継番組を全自動で制作し配信する「ロボット実況」サービスを一般向けに試験提供した(NHK, 2018). また、人間の実況(・解説)がついた中継番組に対し、画面を見ていれば自明であるために言及されにくい情報を選び、同じ枠組で音声化したものを番組とあわせて聴取すれば、現状制作が難しいとされているスポーツ番組の視覚障がい者向け解説放送の代用として機能する可能性がある. 我々はこれを「自動解説放送」と呼び、放送音声との併聴が容易になる自動解説音声挿入タイミングの自動決定手法の研究などを進めている(一木ほか, 2018).

2. 自動実況

試合の今の状況を表すライブデータは試合進行に伴って刻々と変化する. 一方、聞き手はこれまでの一連の実況発話から試合の今の状況に関する信念を構築しており、自動実況生成は、発話タイミングの都度都度において、データと聞き手が持っているであろう信念の差異を解消するための発話を生成し、音声化して伝達する処理を基本とする. 本手法における聞き手信念モデルは、発話によって伝達されたライブデータ中の情報は、その情報の種類に応じて一定の期間記憶されるというものである. この枠組を利用することで、冗長性を抑えて、結束性のある、試合状況への追従性を高める発話を生成(例えば、選手のフルネームはたまにしか言わない、など)したり、状況の変化がなくとも一定間隔で状況のサマ리를発話(いわゆる「堰き止め」)したりすることを可能とする. 現状、この仕組みは、データの信念の差異の種類に応じた選択ルール付きの発話テンプレートをあらかじめ人手で記述することで実現している.

実況発話は助詞の省略や体言止めなどを多く含み、その音声による適切な伝達のためには韻律情報が重要である. ロボット実況試験サービスにおいては、この実現のために新たに音声合成器を開発して利用した.

3. 自動解説放送

テレビ中継の実況では、対戦カードや現在の得点など、画面に表示されている情報への言及が少なめで、音声のみを聴取しても試合状況が把握しづらいことが多い. テレビ中継はラジオに比べて数多く提供されており、視覚障がい者が健常者と同様にスポーツを楽しむためには、このようにテレビ中継を活用した取り組みが重要と考える.

会場音声、実況者、解説者、自動解説音声のすべてを使って試合を楽しむためには、新たに加わる自動解説音声によりよくコンテンツに溶け込むことが大切である. 我々はこれまでに、声質や挿入タイミング、別スピーカーによる提示など、いくつかの観点から調査を行った. 引き続き検討を進めていきたい.

4. おわりに

自動実況、自動解説放送、いずれも、客観データに特化した伝達という点では、従来の実況や解説放送とは機能が異なるもので、その効用やよりよいサービスのあり方についての検討はまだ十分とは言えない. 幅広い観点からご意見をいただければ幸いである.

参考文献

熊野正 (2017). スポーツ番組を解説する「音声ガイド」生成技術 NHK 技研 R&D, **164**, 49-55.

<https://www.nhk.or.jp/str1/publica/rd/rd164/pdf/P49-55.pdf>

NHK (2018). (報道資料) ピョンチャン五輪のデジタル展開の実施結果について.

<https://www.nhk.or.jp/pr/keiei/shiryoku/soukyoku/2018/03/002.pdf>

一木麻乃・熊野正・今井篤・都木徹 (2018). 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 ～ スポーツ中継における実況音声の発話未予測 ～ 信学技報, **118**(270), 45-50.