

聴覚障がい者を対象とした競技場の音の認識と可視化

矢島義久[†] 細井健太[†] 西田昌史[†] 綱川隆司[†] 西村雅史[†]静岡大学 情報学部[†]

1. はじめに

聴覚障がい者の音支援として振動による報知音の伝達、生活音の認識に関する研究がある^[1,2]。どちらも日常生活を送る上で聴覚障がい者に伝えなければならない重要な音を伝えることが目的である。しかしこのような研究は日常生活の範囲にとどまっており、スポーツという分野ではあまり検討されていない。実状として障がい者のためのスポーツを観戦する環境が未だに整っていない。特に聴覚障がい者に向けたサービスが少ない^[3]。聴覚障がい者はアナウンスなどの音声聞こえないので、試合の詳しい状況がわからない。また周りの声援・歓声や応援歌が聞こえないので一体感を味わうことができない。

この課題に対処するため、我々は競技場の音の可視化を目指している^[4]。先行研究では競技場の音の可視化を行うに当たり、競技場の音の認識について検討を行った。先行研究ではオフラインで3種類のスポーツを対象に重なりがないイベントにおいては高い精度で認識できる可能性が示唆された。本研究では実用性を考慮して競技場の音の認識を実時間動作可能な手法について検討を行った。また、SmartEyeglassを用いて競技場の音を可視化する方法についても検討を行ったので報告する。

2. 実時間動作可能な競技場の音の認識

対象としたスポーツはサッカーとバスケットボールである。それぞれの具体的なイベントの種類をTable1に示す。実際に競技場の観客席でスマートフォンの録音アプリを用いて収録した。サンプリング周波数は44.1 kHzである。音データが入力されると、0.5秒ごとに音のデータを分割し、特徴抽出を行い、フレーム単位で識別を行う。機械学習で用いる識別器のモデルとしては、Support Vector Machine (SVM)を利用し、フレーム単位の識別を行った。パラメ

Stadium Sound Recognition and Visualization for Hearing Impaired People.

[†]Yoshihisa Yajima, Kenta Hosoi, Masafumi Nishida, Takashi Tsunakawa, Masafumi Nishimura, Faculty of Informatics Shizuoka University.

Table 1 Event type of each sport

競技名	イベント名
サッカー	歓声
	ホイッスル
	応援歌
バスケットボール	歓声
	ホイッスル
	応援歌
	ブザー

ータは変更せず、カーネル関数はRBFカーネル(Gaussianカーネル)を使用した。あらかじめ学習データによりモデルを学習しておく。特徴量は環境音認識でよく使われているMel-frequency cepstral coefficients (MFCC), Perceptual Linear Predictive (PLP)を使用した。フレーム長を100 ms, フレームシフト幅は10 msとし、それぞれ24次元の特徴量を抽出した。

3. 評価実験

学習データとしてサッカーは各イベント20秒程度、合計90秒程度のものを使用した。バスケットボールは各イベント25秒程度のもの、合計130秒程度のものを使用し、実験を行った。

テストデータとして各イベントをそれぞれつなぎ合わせた120秒程度のものを各競技に3種類用意した。テストデータの識別の結果をTable.2に示す。歓声+応援歌は歓声と応援歌が重なっている音、応援歌+ホイッスルは応援歌とホイッスルが重なっている音、ブザー+歓声はブザーと歓声が重なっている音である。なお、イベントが重なっている区間はどちらかのイベントに認識されれば正解とした。サッカーに比べバスケットボールの正解率が高いのは屋内競技のために雑音が少ないためだと考えられる。MFCCとPLPを比べるとイベントごとに精度の違いが見受けられた。歓声や応援歌などほかのイベントと重なっていない人の声がかかわる部分ではMFCCの精度が高くなっている。しかし重なっているイベントであるとPLPが高い

Table 2 Recognition result of each sport

競技名	イベント名	認識率 (%)	
		MFCC	PLP
サッカー	歓声	90.8	88.6
	ホイッスル	84.4	85.3
	応援歌	78.9	73.5
	歓声+応援歌	82.0	84.9
	応援歌+ホイッスル	91.5	91.9
バスケットボール	歓声	97.2	96.4
	ホイッスル	95.5	95.6
	応援歌	91.2	91.0
	ブザー	93.7	96.2
	ブザー+歓声	96.3	95.5
	歓声+ホイッスル	86.5	87.6

精度になっている。

以上の結果から、実時間動作可能な手法でも高い精度で競技場の音を認識できることがわかった。

4. 競技場の音の可視化

音情報の提供手段に関しては、水野の調査よりスマートフォンや競技場のディスプレイを活用する意見があったが^[3]、本研究ではFigure1に示すSmartEyeglassを使用することで、スマートフォンアプリで実装するのと比べ視点移動を少なくし、より自然に近く観戦に集中しやすい環境の構築を行った。

現在研究を進めている競技場の音の可視化システムでは、SmartEyeglass上に文字情報として歓声、ホイッスルやブザーといった音の表示を行う。なお、可視化システムでは競技場の音の認識ができることを前提としている。

試作した可視化システムの表示画面の一部がFigure2である。付属のスマートフォン用エミュレータを使用し、サッカーの試合の動画を背後に映し録画したものである。象徴的な図形を表示することで、表示情報を明解にした。また



Figure1 Sony SmartEyeglass SED-E1



Figure2 Example of stadium sound visualization

音の種別だけではなく、どのような音なのかを知りたいという聴覚障がい者の実際の声から、「ホイッスル=ピッー ピッー」といった擬音語表示を行った。Figure2のホイッスル表示はさらに視覚的にわかりやすいように点滅表示を実装している。さらに同時に複数の音がなっている場合には、文字の大きさを変えることで音の大小を表現した。

5. おわりに

本研究では、2種類のスポーツを対象に実時間動作可能な手法による競技場の音の認識ならびに、SmartEyeglassを用いた競技場の音の可視化について提案した。実験の結果、実時間動作可能な手法においても高い精度で認識できる可能性が示唆された。また競技場の音の可視化システムについての検討を行った。今後、競技場の音の認識について評価データを増加し、他のスポーツでも評価を行う。また競技場の音の可視化システムの評価実験を行いシステムの改善を行う。

参考文献

- [1] 織田修平ら, 聴覚障害者支援を目的とした振動による報知音の伝達方法とその有効性, 電子情報通信学会論文誌, Vol. J89-D No.12 pp.2671-2678, 2006.
- [2] 浅井研哉ら, 聴覚障害者支援のための実環境下における環境音認識システムに関する検討, 情報処理学会研究報告, Vol.2017-AAC-5 No.11 pp.1-6, 2017.
- [3] 水野映子, 誰もがスポーツ観戦を楽しむための情報提供のあり方, 第一生命経済研究所, 2017.
- [4] 矢島義久ら, 聴覚障がい者を対象とした競技場の音の可視化に関する検討, 日本音響学会講演論文集, 2018.