



音声による接触判定システムの 複数同時運用による全身化の検証

峰松丈倅¹⁾, 前田裕作²⁾, 櫻井翔²⁾, 広田光一²⁾, 野嶋琢也²⁾

Takeyuki MINMATSU, Yusaku MAEDA, Sho SAKURAI, Koichi HIROTA, and Takuya NOJIMA

1) 電気通信大学 情報理工学域 I 類

2) 電気通信大学 大学院情報理工学研究科情報学専攻 (〒182-0035 東京都調布市調布ヶ丘 1-5-1
{minematsu_takeyuki, maeda_yusaku, sho, hirota}@vogue.is.ucc.ac.jp, nojima@nojilab.org)

概要: 我々はこれまで、Augmented Sports 化されたチャンバラ競技, Parablade での利用に向けた、音声認識を用いた剣の接触判定技術を開発してきた。音声を用いることで、少ない装置で広範囲の接触を検出できるようになることを見込んでいる。これまで単一のマイクユニットを試作し、認識精度として 80-90%程度を達成可能であることを確認した。本報告ではこのユニットを全身に複数配置することによる認識への影響を検証し、認識精度の向上にむけた対策を実施したのでその結果について報告する。

キーワード: 音声認識, ウェアラブルデバイス, Augmented Sports, Parablade

1. はじめに

スポーツにおける身体能力格差は、モチベーションの低下やドロップアウトといった悪影響を起こしうる。この悪影響の軽減を行うことを目的に、Augmented Sports [1] が提案されてきた。Augmented Sports は既存のスポーツをベースとし、そのプレイヤーに HP や攻撃力といったバーチャルパラメータ(VP)を付与する拡張したスポーツの形態である。VP の付与により各プレイヤーの身体はバーチャルに能力拡張が適用され、格差が補完されることにより、生来の身体能力や体格差によらずスポーツを楽しめるようにする取り組みである。この Augmented Sports の概念をチャンバラに対して適応し、拡張したスポーツとして我々は Parablade を提案してきた[2]。この競技では VP ならびに打撃身体部位の違いによるダメージ量調整を導入することにより、身体的な能力差を補完し、より多くの人が楽しめるスポーツにすることを目指している。

Augmented Sports の実現においては、まず VP によるダメージの与え方に関するゲーム設計が必要となる。さらに、身体になにかが接触するなどの物理世界のイベントを、VP のダメージ計算に反映させるためのリアルタイム計測の 2 点が大きな課題となる。本稿ではとくに

Parablade における、「剣が相手身体のどの部位にあたったかどうか」を判定するためのシステムについて述べる。

Parablade では試合中に何度も打撃が発生し、打撃ごとに与えるダメージが異なる。現状の Parablade ゲーム設計で

は、剣が相手の腕と足どちらにあたったかにより、ダメージ量が異なる設計となっている[3]。そのため、打撃の有無や接触部位を高速に計測・判定可能なシステムが必要となる。

この課題に対して我々は、小型マイクを用いた音声認識による接触部位判定システムを開発してきた。音声を用いることで、例えば圧力センサなどと異なり、センサの存在する部位以外の衝撃も判定可能となり、より少ないセンサでより広い範囲の剣と身体の接触判定が実現可能になると期待している。使用する剣がプレイヤーに衝突する際の多様な音声を記録・学習することで、マイク近傍での接触の発生有無判定を試みるものである。これまでの試作の結果、小型マイク 2 個からなる単一ユニットで 80-90%ほどの認識精度を出すことができた。しかし単一ユニットでは全身の剣との接触を判定するには不十分であり、またそもそもスポーツにおける判定システムとしては十分な精度とは言い難い。そこでまずこれらの課題のうち、複数のマイクユニットを全身に複数配置することによる全身と剣との接触判定への取り組み、およびその接触判定精度への影響について検証を行ったので、詳細を報告する。

2. 先行研究

剣術競技における接触判定技術として、フェンシングでの利用が代表的事例として挙げられる。プレイヤーはラ

メと呼ばれる金属製素材により構成されるベストやグローブを装着し、そこに相手の保持する剣が接触したときに両者の間に電流が流れることで接触判定を行っている[4]。この手法では接触有無の判定に多くの実績はあるものの、接触部位の判定は試みられていない。技術的には可能であるが、身体全体の接触部位判定の実現にはシステムの複雑化を招き、プレイヤーの動きを制限することに繋がると考えられる。よりシンプルな取り組みとして圧力センサによる接触判定がある。SASSEN では剣に圧力センサを内蔵し、その反応によって接触の判定を行っている[5]。しかしこの判定方法の場合、そもそも接触物が身体であるかどうかを判定することはできず、またその接触部位の同定も困難である。接触の有無の判定という観点では、人体通信技術も多くの活用事例がある[6]。しかしこれも接触部位の判定が困難であるという課題を抱えている。

これまで身体への接触の有無を判定する技術について多くの取り組みがなされてきたものの、スポーツへの利用を前提としたうえで、身体の中のどの部位に接触したか、という部位判定まで踏み込んだものは多くない。そこで本研究では小型マイクユニットを用いて剣の衝突音を収集判定することによる、接触部位同定可能な接触判定技術の開発を実施している。デバイスが小型であることから、プレイヤーの動きを制限することなく、少数のデバイスで広範囲の接触有無ならびに接触部位判定の実現を目指すしている。

3. Parablade の接触有無・部位判定技術

3.1 Parablade の概要

Parablade はチャンバラ競技を Augmented Sports 化した競技である。チャンバラ競技のプレイヤーに、HP (体力) /ATK (攻撃力) /DEF (防御力) /AGI (攻撃頻度) の 4 つからなるバーチャルパラメータ (VP) を導入することで実現されている[3]。これら VP を適切に設定することにより、物理的な身体格差の影響を軽減し、誰もが楽しめるスポーツの実現を目指している。

Parablade ではその設計理念上、打撃が身体の中のどの部位に命中したか、打撃部位によってダメージ量に変化する設計が導入されている。一撃で大きなダメージを狙う、当てやすい代わりに小さなダメージを多数与える、など、戦略の多様化を狙って導入された仕組みである。しかし、これを実現し Parablade をストレスなくプレイするためには、打撃の命中部位を高速かつ高精度で判断できるシステムの実現が必要となる。そこで Parablade では剣の内部に圧力センサが組み込まれている SASSEN 刀を利用するものとし、SASSEN 刀と身体との接触時に発生する音声を学習し、判定に活用することで、この問題の解決に取り組んできた。

3.2 20kHz 超音波補助音源を用いた接触判定システム

我々はこれまで、SASSEN 刀が身体に衝突する際の音声特徴を記録・学習することで、音声による接触判定に取り組んできた。その際、SASSEN 刀にスピーカを搭載し、補助音源として 10kHz の音声を SASSEN 刀より出力することで、判定精度の向上に努めてきた。しかし 10kHz の音声は可聴域でありプレイの妨げになること、また当初開発した録音用のマイクユニットが長さ 17cm 程度と比較的大きいことから、その小型化に取り組むとともに、補助音源の周波数を可聴域限界領域の 20 kHz とするシステムの構築を行った。超音波補助音源装置 UTA(Ultrasound Tsuba Attachment)は図 1 のように刀の鍔を模した構造で作られており、SASSEN 刀の持ち手と刀身の境界部に装着される。一方身体側には、図 2 のようなマイクユニット R-PMU (Recording Parablade Microphone Unit)を搭載する。

本研究では接触判定システムの基礎を構築するため、1 器の R-PMU を左腕に装着し、R-PMU 近傍を含む全身を UTA 装着済みの SASSEN 刀でたたき、その音声を収集する。収集した音声を学習することで、音声情報から UTA 装着済み SASSEN 刀が叩いた身体部位の判定を試みる。R-PMU 上のマイクより得られた音声はマルチトラックレコーダを通してサンプリング周波数 44.1kHz、サンプルサイズ 16bit で録音される。この R-PMU で収集した音声から打撃音を含む 1.0 s の区間を抽出し、この音声に窓枠 30ms、窓間隔 20ms でハン窓を用いた短時間フーリエ変換(STFT)を行うことで、時間方向には 49 フレーム、周波数方向には学習のためにメルバンクフィルタなどでの圧縮を行い、ビン数 40 とした 2 次元スペクトログラムを作成する。このスペクトログラムを 2 次元畳み込み層(8 filters, 10x1 kernel, 2x2 stride, depth 8, bias, ReLU activation) 1 層、全結合層 (4 units, weights 4x4000, bias), 及び出力層 (4 units, beta 1) から構成される 2 次元 CNN モデルで学習した分類精度は図 3 のようになった。この図 3 ではある音声での実際の命中部位が行で、モデルが判定した命中部位が列で示されており、各セルには実際の命中部位と同じ音声のうち、モデルがその部位だと判定した百分率を記載している。つまり、対角線上のセルに記載されている百分率が、モデルの分類精度となっている。そしてその精度を見ると全体としては 8 割程度の精度であり、またユニットの装着部位である腕への打撃の精度が 88.2%とほかの部位よりも高いことが確認できた。この精度は目標である 95%に満たないものであり十分とは言えない。そのためユニットの装着部位に近い位置の打撃は精度が高く、これを利用すれば他の部位の精度向上も狙えるのではないかと考え、4 章のようにユニット数を増加させた状態での検証を行うこととした。

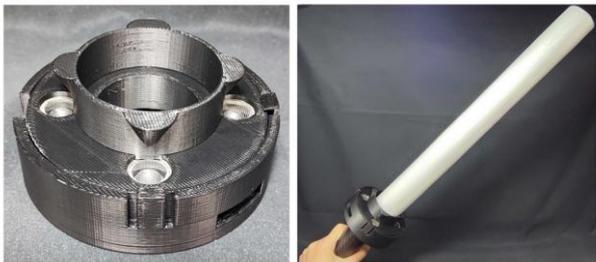


図 1:補助音源装置 UTA とそれを装着した SASSEN 刀

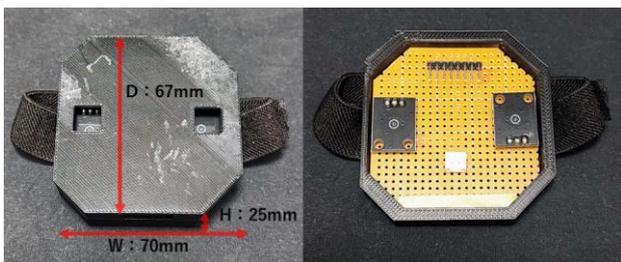


図 2 : マイクユニット R-PMU

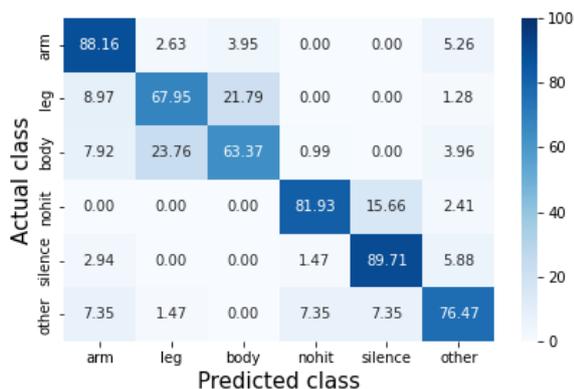


図 3 : 単一ユニットでの分類精度

4. 4 ユニット構成による判定システム全身化

4.1 音声の録音学習, および分類精度の検証

これまで判定部位は arm, leg, body であったが競技中に剣の突きの動作が胴体に命中することでけがのリスクがあることが判明したため、body を判定対象から外し、代わりに arm を leftarm と rightarm, leg を leftleg と rightleg と分けて判定を行うこととした。本実験で用いた音声は図 4 のように身体の 4 箇所につけた状態でその他の録音条件は 3.2 と同様にして 4 ユニット同時に録音を行うことで学習に用いるデータセットを用意した。この条件において、3.2 と同様の学習を実施し、各ユニットそれぞれについて個別の判定モデルを構築した。その判定精度の検証結果を図 5 に示す。

3 章の結果とは分類項目が異なるため、直接的な比較はできないものの、ユニット装着部位への打撃に対する精度が高い傾向があることが確認でき、判定精度は概ね同程度を維持できていると考えられる。しかしそれでも、低いときには 70% 台の判定精度であり、十分とは言

い難い。またこの手法の場合、接触判定が 4 つ出力されるため、ここから単一の接触イベントの発生について、さらなる検証が必要になる。



図 4: 録音時の R-PMU の配置

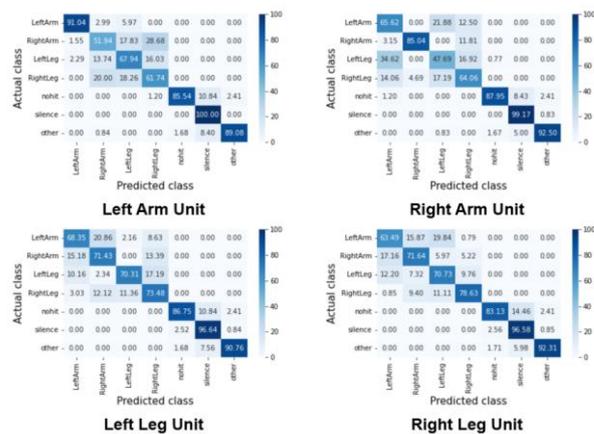


図 5: 各ユニットでの分類精度

4.2 複数ユニットデータを用いた単一判定モデル学習

この節では 4 ユニットの音声を同時に学習することで 1 つのモデルを作成し、それによる情報量の増加により判定精度向上を狙った結果について報告する。

この学習ではまず 4ch の音声を図 6 のような前処理を行った上でモデルに入力した。同時に入力するスペクトログラムには共通の時間シフトと背景ノイズの追加を行った。モデルではこの入力を 4 チャンネル入力ができるようにテンソルの形を再構成した上でこのデータを 4 チャンネルの畳み込み層に入力することで 4 ユニットの同時学習を可能とした。この学習モデルは入力チャンネル数を 1 から 4 に変更した以外は 3.2 のものと同様であった。この学習モデルでの分類精度は以下の図 7 のようになった。

この結果を図 5 の各ユニットの結果と比較すると各ユニットの装着位置への打撃の判定精度に対しては大きく劣っていることが確認できる。例えば左腕ユニットにおける左腕の精度 91.0% と比べるとこのモデルでの左腕の 70.9% と精度が低いものになっている。このため、このモ

デルは4ユニットの音声情報を十分に生かしているとは言い難いものであった。

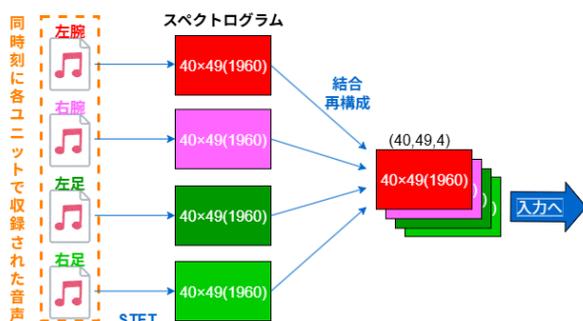


図 6:4ch 入力の前処理

Actual class	LeftArm	RightArm	LeftLeg	RightLeg	nohit	silence	other
LeftArm	70.90	17.16	10.45	0.00	0.00	0.75	0.75
RightArm	20.16	64.34	4.65	10.08	0.00	0.00	0.78
LeftLeg	20.61	8.40	58.78	12.21	0.00	0.00	0.00
RightLeg	2.61	13.04	10.43	73.91	0.00	0.00	0.00
nohit	1.20	0.00	0.00	0.00	77.11	19.28	2.41
silence	0.00	0.00	0.00	0.00	1.40	95.10	3.50
other	0.00	0.00	0.00	2.46	3.28	5.74	88.52
	LeftArm	RightArm	LeftLeg	RightLeg	nohit	silence	other
	Predicted class						

図 7:4ch 同時学習での分類精度

この原因の一つとして、今回構築したモデルは畳み込み層が一層しかなく、かつプリーング層を省いたものになっていることの影響が考えられる。この学習に関する一連の実験では、R-PMU で音声を集積し、そのデータを PC 上で学習しモデル構築を行っている。しかし本来、R-PMU はマイコンを搭載し、そこに判定モデルをインストールして、単体で判定を実施できるようにすることを考えている。そのため学習モデルも、将来的なマイコン搭載を睨んで簡略化されている。しかしこの簡略化により過学習に弱く、それによる精度の低下起きた可能性や畳み込みによる各チャンネル間の融合ができておらず、モデルが各チャンネルの位置関係を十分に生かせなかった可能性が考えられる。そのためマイコン搭載が可能な範囲での畳み込み層の増加や 3D 畳み込み層の利用、全結合層を GAP 層に置き換えるなどの対策を行うことでの精度向上を試みる。

5. まとめ

本報告では、従来提案していた音声による接触判定システムについて、判定補助音源の周波数を変更したうえで、左右の手足合計 4 箇所を搭載し、全身での接触判定の実現にむけた基礎検討を実施した。各ユニット個別の判定精度は 70~90%程度を達成可能であるものの、その

プレイヤーに剣が接触したかどうか、どの部位に接触したか、という全身での接触判定については 50~70%程度と振るわず、さらなる精度向上が必要であるという結果を得た。また、検証の過程で判定のリアルタイム性にも課題があることが判明した。現状では打撃から判定までに 1.1s ほどの時間がかかり、プレイヤーへのストレス、試合進行への悪影響が懸念される。

これらの課題を踏まえて今後は、学習モデルの最適化による判定精度の向上を狙うとともに、学習モデルに利用する音声の時間枠を短縮化することによるリアルタイム性の向上に取り組む。これらを統合してシステムの安定性と認識精度の向上を達成し、試合で実用できるような Parablade の部位接触判定システムの完成を目指す。

6. 謝辞

本研究の遂行にあたって、全日本 SASSEN 協会の本村隆馬氏、鋤先星汰氏に多大なご協力をいただきました。この場を借りて御礼申し上げます。

参考文献

- [1] Kadri Rebane, Takahiro Kai, Naoki Endo, Tomonari Imai, Takuya Nojima, Yohei Yanase. Insights of the augmented dodgeball game design and play test. AH '17: Proceedings of the 8th Augmented Human International Conference. Article No.12. Pages 1–10. 2017.
- [2] 前田 裕作, 佐藤 神威, 櫻井 翔, 広田 光一, 野嶋 琢也. マイクを利用した模擬刀と身体の接触判定システムの提案, 第 28 回日本バーチャルリアリティ学会大会論文集, 1C1-09, 2023.
- [3] Naoto Nishida, Yusaku Maeda, Kamui Sato, Sho Sakurai, Koichi Hirota and Takuya Nojima. Parablade: A Proposal of Chambara Based on Augmented Sports - A Study of Appropriate Game Balancing Methods. The 24th IFIP INTERNATIONAL CONFERENCE ON ENTERTAINMENT COMPUTING. 2025. (TBA)
- [4] Craig Freudenrich. Ph.D. How Fencing Equipment Works. <https://entertainment.howstuffworks.com/fencing-equipment.htm#pt3> (Visited 2025/7/16)
- [5] 一般社団法人 全日本 SASSEN 協会. 次世代のデジタルスポーツコンテンツ SASSEN (サッセン) とは, <https://sassen.jp/about/#03015> (Visited 2025/07/11)
- [6] 蜂須拓, 鈴木健嗣. 身体の物理的接触を拡張するウェアラブルデバイス (第 9 報): 電流式人体通信のための電極最適化基礎検討, 第 25 回日本バーチャルリアリティ学会大会論文集, 1C1-1, 2020.