



プレイヤー自身の感情制御を余儀なくされる 対戦型ゲームの開発

千葉工業大学情報科学部情報工学科
准教授 有本 泰子

1. はじめに

ヘッドマウントディスプレイなどで VR を利用したゲーム開発が人気を集める中、ゲームに対する没入感を向上させ、よりリアルなバーチャル空間を実現することが求められている。本研究では、対戦型ゲームのプレイ中に無意識に表出する感情を入力とし、アバターの状態をプレイヤーの意図に関わらず強制的に変更（ヒットポイントの増減や対戦相手への攻撃、場面転換など）する「感情入力インタフェース」の開発を目指す。本システムにおける感情表出は音声に含まれる Affect Burst を対象とする。Affect Burst とは音声の非言語的使用による感情の急激な表出のことで、無意識な叫び声や笑い声、泣き声や、「ゲッ」「おほ！」など言語音として意識的な制御下にあるものを含む[1]。本システムは、ゲーム中の笑い声や叫び声を検出・識別する Affect Burst 検出部と、検出された Affect Burst の種類をもとめてゲーム内の状態を変化させるリアクション生成部の二つのモジュールで構成される。

本研究課題では、Affect Burst 検出実験と、主観的評価実験によりプレイヤーの没入感を向上させるリアクションについて検討した。本報告書ではその成果についてまとめる。

2. Affect Burst 検出実験

プレイヤーが思わず発する Affect Burst を、その音響的特徴から機械学習によって検出する仕組みを構築する。本研究で使用する笑い声および叫び声は、ゲーム中のユーザのリアルな対話を収録した二つのコーパスを使用している[2][3]。

2.1. 叫び声識別実験

叫び声検出に向けて、まずは人手により叫び声区間とそれ以外の発話区間を区切ったデータを機械学習により識別する実験を行った。使用したコーパスに含まれる発話数は14,124 発話で、そのうち叫び声の数 1,427 発話である。openSMILE[4]の Interspeech 2013 ComParE Challenge baseline feature set を使って、基本周波数・MFCC の一次係数・音圧レベルなどを含む 6,373 次元の特徴量を抽出した。識別器には 5 層の隠れ層 (2048 ユニット) とそれに続く Dropout 層を持つディープニュー

表 1 叫び声の識別結果

	正解率	適合率	再現率	F-measure
学習用	97.7%	87.1%	90.3%	88.7%
検証用	96.7%	84.5%	82.8%	83.7%
評価用	96.0%	80.4%	82.2%	81.3%

表 2 笑い声の検出結果

	適合率	再現率	F-measure
笑い声	79%	74%	76%
発話	91%	91%	91%
無音	97%	97%	97%

ラルネットワーク (DNN) を用いた。識別の結果、学習用・検証用・評価用の各データに対する正解率・適合率・再現率・F-measure を表 1 に示す。評価用データの再現率が 82.2%であることから、叫び声はある程度の精度で発話と識別可能となることが示された[5]。

2.2. 笑い声検出実験

対話中にいつ笑い声が出現したかを検出するモデルを作成した。対話音声を 1 秒ごとに笑い声区間・発話区間・無音区間と区切ってラベルを付与し、叫び声と同じく openSMILE を用いた 6,373 次元の音響的特徴量を抽出した。笑い声の比率は全体の 6.07% である。検出器には 6 層の隠れ層 (2048 ユニット) とそれに続く Dropout 層 (Dropout 率は 0.2) を持つ DNN を用いた。検出の結果、笑い声・発話・無音の各区間に対する適合率・再現率・F-measure を表 2 に示す。笑い声の再現率は 74% であることから、ある程度の笑い声検出は可能であるものの改善の余地があることが示された[6]。

3. 笑い声に適切なリアクションの選定

プレイヤーが思わず発した叫び声や笑い声に対してどのようなリアクションを返すことが適切か決定する必要がある。しかしながら、どのようなリアクションを返すとプレイヤーがよりゲームに没入するか明らかになっていない。そのため、まずはプレイヤーにとって有利 (ポジティブ) なリアクションを返すべきか、不利 (ネガティブ) なリアクションを返すべきかという大きな枠組みで検証実験を行った。

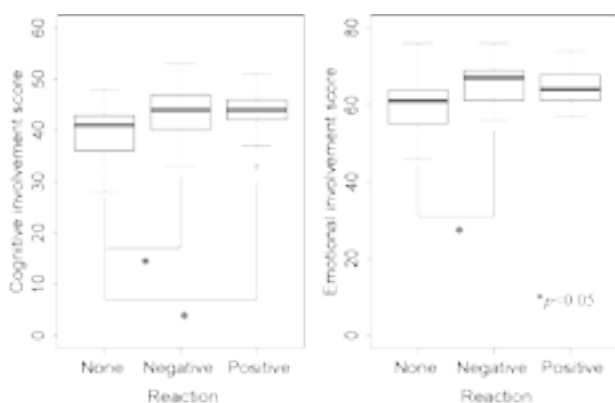


図 1 主観評価実験の結果

実験に利用するゲームを Minecraft[7]に限定し、Modification (MOD) 開発を行って、いくつかのイベントを利用可能にした。さらに、実装したリアクションの種類を、笑い声を発したプレイヤーにとってポジティブなリアクションとネガティブなリアクションに分類した。

被験者 12 名にこれらのリアクションを返すゲームプレイを行わせた。各被験者にはリアクションなし・ネガティブリアクション有り・ポジティブリアクション有りのすべての条件でゲームプレイさせ、それぞれの条件に対してどの程度没入したかを計測する主観評価を行わせた。各条件の没入感を評価するために使用した質問紙はゲームの没入感を評価するために開発された Immersive Experience Questionnaire (IEQ) [8]である。IEQ は 5 因子から構成されており、積極的に関与と注意を測定する Cognitive Involvement、感情と不安を測定する Emotional Involvement、周囲や現実世界にいるという認識の欠如を測定する Real World Dissociation、ユーザがゲームにこれだけ困難を感じたか測定する Challenge、ゲームインターフェースの使い易さを測定する Control で構成されている。

主観評価実験の結果の一部を図 1 に示す。図 1 にはリアクションなし (None) ・ネガティブリアクション有り (Negative) ・ポジティブリアクション有り (Positive) の条件ごとに、Cognitive Involvement と Emotional Involvement の各因子のスコアを箱ひげ図で示している。プレイヤーの笑い声に対してネガティブなリアクションを返した方が、リアクションが全くないよりも認知的にも感情的にもゲームに引き付けられることが示された。また、ポジティブなリアクションを返した場合は、認知的にゲームに引き込まれるものの、感情的には引き付けられないことが明らかになった[9]。

Cognitive Involvement のスコアがリアクションの種類にかかわらずリアクションなし条件より高くなったのは、笑い声に対して何らかのリアクションを発生させると無意識に注意がゲームに

向けられ、積極的に関与してゲームをコントロールしようとした結果であることが考えられる。一方で、Emotional Involvement のスコアがネガティブリアクションを提示したときに高くなった理由は、プレイヤーにとって不利なリアクションが思いがけず発生したため、不快な感情が強く表れた結果であると考えられる。

4. まとめ

無意識に表出する感情を入力とし、その内容によってアバターの状態をプレイヤーの意図に関わらず自動的に変更する「感情入力インタフェース」の開発に向けた検証を行った。その結果、叫び声は 82.2% の精度で識別が、笑い声は 74% の精度で検出が可能であることを示した。また、プレイヤーが思わず笑ってしまったときに、よりゲームに没入させるリアクションはプレイヤーにとって不利になるリアクションであることが分かった。

参考文献

- [1] M. Schröder, "Experimental study of affect bursts," *Speech Communication*, vol.40, no.1-2, pp.99-116, 2003.
- [2] 菊地右樹, 森大毅, "音声コミュニケーションにおける叫び声," *日本音響学会 2017 年秋季研究発表会講演論文集*, pp.399-400, 2017.
- [3] Y. Arimoto, H. Kawatsu, S. Ohno, and H. Iida, "Naturalistic emotional speech collection paradigm with online game and its psychological and acoustical assessment," *Acoustical Science and Technology*, vol.33, no.6, pp.359-369, 2012.
- [4] F. Eyben, M. Woellmer, B. Schuller, openSMILE the Munich open Speech and Music Interpretation by Large Space Extraction toolkit (2010).
- [5] 金子裕亮, 有本泰子, "音響的特徴量によるゲームプレイ中の叫び声の識別!" *日本音響学会 2020 年春季研究発表会講演論文集*, pp. 995-996, 2020.
- [6] クー ユエン, 有本泰子, "ゲームプレイ中の笑い声検出モデルの構築" *日本音響学会 2020 年春季研究発表会講演論文集*, pp. 1005-1006, 2020.
- [7] Mojang, Minecraft, <https://www.minecraft.net/ja-jp/>
- [8] C. I. Jennett, "Is Game Immersion Just Another Form of Selective Attention? An Empirical Investigation of Real World Dissociation in Computer Game Immersion," *Doctoral thesis, University College London*, 2010.
- [9] 深津多聞, "無意識な笑い声にリアクションするゲームの開発に向けたゲームイベントの実装とその評価," *帝京大学学士論文*, 2020.