

仮想対戦プレイヤーの感情的発話生成

塩入健太 星野准一

筑波大学大学院 システム情報工学研究科

概要 対戦型コンピュータゲームをプレイヤー1人でプレイする場合、人間のプレイヤー相手にプレイする場合と比較してつまらないと感じるプレイヤーが多い。その主な理由のひとつに「対戦相手の AI の反応が乏しい」といった問題がある。その改善のために本稿では、ゲーム映像からゲーム状況を読み取り、プレイヤーからの干渉や環境変化に対して発話音声によって反応を返す「仮想対戦プレイヤー」を提案する。またそのとき、感情的な内容や声の調子、発話頻度の出力をするために2つの内部パラメータ値を参照して音声ファイル選択を行う。これらはプレイ中にゲーム状況によって動的に変化させ、現在に至るまでの心理状態変化の蓄積として扱う。最後に手法の有意性を評価実験により確かめる。

Generating Emotional Utterance of the Virtual Adversary Player

Kenta Shioiri Junichi Hoshino

University of Tsukuba Graduate School of System and Information Engineering

Abstract-A player feels bored when he play the match-up type computer game without adversary person. One of the main reasons is "The reaction of adversary AI is scarce". To improve this problem, this paper proposes "The Virtual Adversary Player" that reads the game situation from game image and returns the reaction by the utterance voice when the player interferes or the environmental is changed. Moreover, in order to output emotional content, voice condition and the utterance frequency, the voice file selection is done referring to two internal parameter values that dynamically change by the game situation while playing and are treated as accumulation of the psychological condition change. At the last, the evaluation experiment of the technique is conducted.

1. はじめに

コンピュータゲームの誕生以来、カードゲーム、ボードゲーム、スポーツなど多くのゲームジャンルにおいて、「対戦」の要素を含むものが出現してきた。また、家庭用テレビゲーム機が広く普及し、「テレビゲーム」がエンタテインメントの主要分野になった近年において、格闘アクション、カーレース、落ち物など、対戦型ゲームのジャンルは多様化し、多くのタイトルが発売され、人気は衰えない。

対戦型ゲームをプレイする際、部屋にプレイヤーが1人のみの状況、またはインターネットを用いての通信対戦や音声チャットなどの環境が整っていない環境では、プレイヤーは通称 COM と呼ばれるゲーム内の AI と対戦することになる。

この COM との対戦であるが、人間との対戦と比較してつまらないと感じるプレイヤーが大半である。その主な理由の1つに「こちらの行動に対して反応が乏しい」といったものがある。

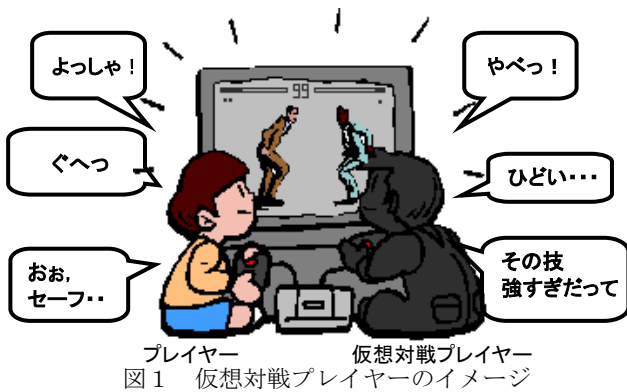
一方、実際の人間が対戦相手の場合には、特に発話によって、対戦中に様々な反応、そして感情を受ける。例えば、危機的状況から脱出した際、対戦相手に攻撃が成功

した際には「よしっ」のような喜びの声を漏らし、逆に困難な状況に陥った場合には「うわあ」、「勘弁してくれよ」といった驚きや非難の声を発す。この多様な反応と感情伝達が対戦型ゲームの面白さのひとつの要因と考えられる。従来のテレビゲームには不利な状況になるとゲームキャラクターの表情が変化したり、汗を出したりするようなものも見られるが、単純で機械的であり、人間と対戦しているような感覚を得るのは難しい。

また近年では、人間からの干渉や周囲の環境変化に対して感情的な反応を返すシステムの開発が盛んに行われており、AIBO[1]やけんかデモ[2]は代表的な例である。

ゲーム AI の研究では、プレイヤーの模倣によって人間らしい反応を行動で返すといったものがあるが[3]、対戦相手からの感情的な反応である発話に関しては触れていない。また CG エージェントやロボットの間人らしい反応に関する研究としては[4][5]があるが、対戦型ゲームという環境への適応は考えていない。

そこで本研究では、対戦型ゲームの 1 ジャンルである「格闘アクションゲーム」を題材に、対戦ゲームプレイ中に



プレイヤー 仮想対戦プレイヤー
図1 仮想対戦プレイヤーのイメージ

プレイヤーの行動やゲーム状況を読み取り、対戦相手としての感情的な発話によって反応を返す「仮想対戦プレイヤー」を提案する。仮想対戦プレイヤーはゲーム中の登場キャラクターとは異なり、実世界でプレイヤーと対戦する人間の模倣モデルである(図1)。

2. 手法の概要

仮想対戦プレイヤーは、入力であるゲーム映像からゲーム状況を読み取り、その場面にあった適切な発話音声を出力する。仮想対戦プレイヤーの姿を画面に表示せず発話のみを用いるのは、実際に格闘アクションゲームをプレイ中のプレイヤーは対戦相手の姿を見る頻度が非常に低く、相手の様子を主に声から感じる事、また画面に表示することでゲームへの集中を妨げること[6]が理由である。声から人間らしく感情的な発話音声を選択・出力するために、仮想対戦プレイヤーの内部パラメータとして快-不快パラメータと発話頻度パラメータを設定する。これらにより同じゲーム状況でも、出力は感情を考慮した声の調子や発話内容・頻度に変化する。パラメータは単発的な攻撃の成否や残りライフの差だけではなく、連続的な過程や全体的な勝率を考慮し算出する。

以下処理の流れである。まず、前処理段階は次の流れで行う。

1. 実際に市販の格闘アクションゲームを人間同士で対戦している様子をビデオで撮影し、発話音声の録音をする。録音した音声から発話音声ファイルを作成す

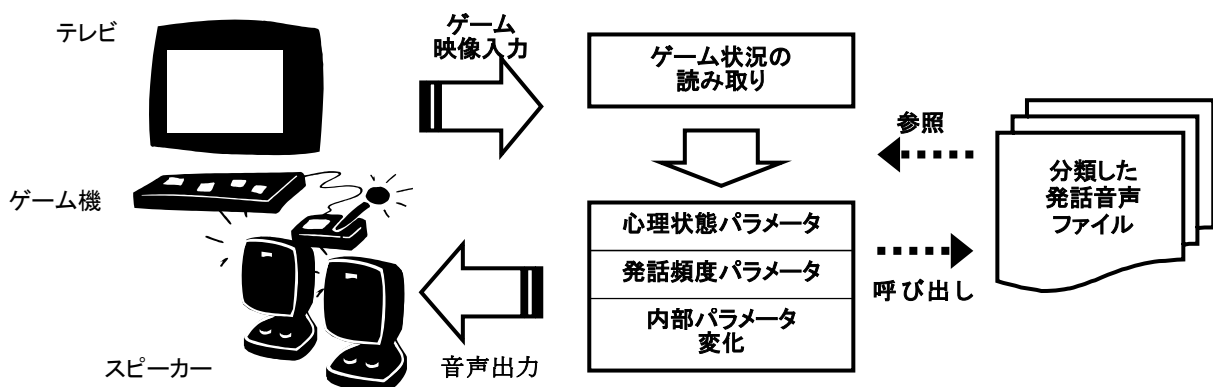


図2 システム概要図

- る。
2. ビデオを参照しつつ、発話音声ファイルを発話発生時のゲーム状況によって4分類、さらに音声から感じるの快-不快の印象によって5分類する。ゲーム状況による分類ごとに発話頻度の初期固有値を設定する。
3. 仮想対戦プレイヤーの内部パラメータとして、心理状態の快-不快の度合いを表す快-不快パラメータと、ゲームへの没入具合に起因し発話頻度を制御する発話頻度パラメータを設定する。ゲーム状況とパラメータ値変化を対応付ける。
4. 快-不快パラメータと発話発生時のゲーム状況を、先ほど分類した発話音声ファイルに対応付ける。

実行段階は次の流れで行う。

5. ゲーム映像をPCに取り込み、その映像から特定のゲーム状況を表す特定の情報を検出し、ゲーム状況を読み取る。
6. 読み取ったゲーム状況により、快-不快パラメータと発話頻度パラメータが変化する。
7. ゲーム状況と、快-不快パラメータを参照して、対応する発話音声ファイルを呼び出す。
8. 発話音声ファイルごとの初期固有頻度と、発話頻度パラメータを参照して、確率により音声出力を行う。

最後に、始めに撮影した格闘アクションゲームに仮想対戦プレイヤーを適応し、被験者がシステムを体験して心理評価を行う。

3. 発話音声ファイルの分類

発話音声ファイルは、実際に格闘アクションゲームをプレイしている最中の特定のプレイヤーの発話を録音し、そこから感情的な発話と感ずるものを抽出することによって作成する。声優によって作られた音声ファイルを用いない理由は、それをを用いた場合の印象評価をした前実験において、ゲームキャラクターの音声と混同してしまい人間と対戦している感覚が薄れると感ずたため、特定のプレイヤーの発話か

らのみファイルを作成する理由は、複数の人物の声を仮想対戦プレイヤーが出力した場合の印象評価を行った前実験において、1体の仮想対戦プレイヤーが複数の人物による声を発するのことに違和感を感じたためである。また特定の人物を起源とする音声ファイルのみを用いた前実験では、その個人と対戦しているような感覚を得ることができた。

実際のゲームプレイ中の様子の撮影と、発話音声の録音は、人物と時間を変えつつ合計5時間前後行った。感情的発話を、感嘆句と自身の状況を表した短文、およびプレイ中のゲームに関係ある感想と定義し、1人の人物が1時間プレイした際には100前後の発話音声ファイルが抽出できた。

この発話音声ファイルを発生時のゲーム内の状況および快-不快によって分類し、仮想対戦プレイヤーが適切なタイミングで適切な発話を行うために用いる。

3. 1. 発生時の状況による分類

撮影したビデオを検証した結果、感情的発話は発生時のゲーム内の状況によって大きく分けて以下の4属性に分類できると考える。

① 予期発話

相手から攻撃を受けることや自分の行動の成功または失敗を予期した際に発せられる。一度に大きなダメージを与えられ、ゲーム画面の明るさが大きく変わったりゲーム内の時間が一時停止したりするような映像上の特殊効果を伴う大技の発生前に多く出現し、単語のみまたは短文の場合が多い。さらに攻撃側と被攻撃側に分類される。

発話頻度の初期固有値は高い。

例・やば、うわっ、危ない、よっしゃ、いける

② 反応発話

相手から攻撃を受けたとき、自分の攻撃が成功したときに反射的に発してしまう感嘆句等。攻撃側と被攻撃側の分類したうえ、さらに予期の有無および攻撃の成功失敗によって3つに分類する(表1)。

予期のある場合は発話頻度の初期固有値が高いが、ない場合は低い。

表1 反応発話の分類(被攻撃側の例)

	攻撃をうけた	回避した
予期あり	反応発話1 うえ、まじでえ	反応発話2 セーフ、助かった
予期なし	反応発話3 うお、む、お	なし

③ 能動発話

攻撃をする際の掛け声。動詞や固有名詞(技名な

ど)であることが多い。ゲーム状況的の観点からは、大技の発生前、心理状態的の観点からは上機嫌で没入しているときに多く出現する。攻撃側に限定される。

発話頻度の初期固有値は低い。

例・くらえ、○○拳、

④ 文章発話

状況についての感想など。他の属性と比較して2文節以上から構成されるような長いフレーズであることが多い。キャラ間が長距離、ダウン中、ライフの変動が一定時間ないなどの試合中の中休み状態、あるいは対戦終了直後に多く出現する。

試合終了時、ライフに大きく差をつけられている状況など、特定された状況での発話と、「(必殺技が)でないっ!」「ミスった」など、ある程度状況に関わらず発せられる言葉がある。特定の状況に起因するものは、この分類全体の単位でゲーム状況に対応付けることは不可能と言える。

発話頻度の初期固有値はファイルごとに異なる。

例・その技強すぎ、このまま何もできず負けてしまうのか、今回は勝てそう、やってみよ、ミスった

その他ゲームに関係する発話には次のようなものもあるが、今回は取り扱わない。

例・○○拳ってどういうコマンドだったっけ?

①②③の発話音声ファイルはその分類単位で、上で説明した出現の傾向に合わせてゲーム場面に対応付ける。

④の文章発話のうち、特定の状況に起因するものは、ファイル1個単位でゲーム場面と対応付ける。状況に関わらず発せられる言葉は、①②③以外の状況でランダムで呼び出す。

3. 2. 音声から感じる快-不快による分類

ゲームプレイ中、プレイヤーは心理状態を動的に変化させる。撮影したビデオを検証した結果、プレイヤーの心理状態に起因していると考えられる内容および声の調子の発話が多く見られた。声の調子は、言葉の意味以上に、心理状態を聞き手に的確に伝える[7][8]。例えば語尾に撥音をつける、語尾を伸ばす、低くゆっくりしゃべる、高く早くしゃべるのでは、それぞれ聞き手が感じる印象が大きく異なる。

本稿では発話音声ファイルを、聴いた際に受ける印象をもとに快-不快の尺度で分類する。これはゲームプレイ中には多種多様で複雑な心理は生まれにくく、心理状態の定量化、そして仮想対戦プレイヤーの心理状態制御には、心理状態の最も原始的な2分類である快-不快で十分だと考えるからである。

具体的には不快、やや不快、普通、やや快、快の5段階のランクに分ける。ここでは不快はイライラ、不機嫌といった負の感情を総合したものであり、快は楽しい、上機嫌といった正の感情を総合したものであと定義する。発話音声ファイルの快-不快のランク分け、および前節の発生時の状況による分類を組み合わせたものの一部を例として以下の表2に示す。

表2 ゲーム状況分類と快-不快分類の組み合わせ

	予期発話	反応発話 (①)	能動発話	文章発話
不快	やめるよ	ちくしょ	しね	最悪～
やや不快	うそお	ええ～	くらえ	勘弁して
普通	まじで	うわっ	いけ	やられた
やや快	きた	くらった	えいつ	うおお
快	お	うまいっ	やー	うまいね

4. 内部パラメータ

本研究では仮想対戦プレイヤーの心理状態の定量化として、初期値 0, 100 から-100 までの定義域を動的に推移する快-不快パラメータを設定し、快から不快までランク分けされた発話音声ファイルの分類と対応付ける。

またそれとは別に、発話発生頻度を制御する発話頻度パラメータを設定する。初期値 50, 0 から 100 の定義域を動的に推移し、数値が高いほど発話頻度が高くなる。快-不快パラメータと発話頻度パラメータを別に設定するのは、連続して勝ち続けており非常に快の状態でも、対戦相手に対する罪悪感などから発話を控える状況、または簡単に勝ち続けることで「飽き」が早まった状況、そして負け続けて不快な状況であっても、イライラして暴言が多くなるといった状況の考慮による。

各パラメータの変動要因は、今回扱う格闘アクションゲームを実際に人間同士でプレイしたプレイヤー数人の意見を参考して決定した。そして各要因ごと簡単な式を設定し、パラメータ値算出に用いる。

4. 1. 快-不快パラメータ値の算出

本稿では、心理状態の快-不快を変化させる要因として、成功した(受けた)攻撃の威力、一定時間内の攻撃(被攻撃)回数、予期した攻撃を外す(回避する)、残りのライフゲージが大きく開いている、試合に勝ち(負け)続ける、を設定し、以下 i ~ v に示す式によって快-不快パラメータ値を算出する。

このときダメージ量や残り時間などの物理量と快-不快パラメータ値の心理量の関係は、諸々の刺激に対して次の関係式が精神物理学的に成り立つというウェーバー・フェヒナーの法則[9][10]に従うと考えられる。

$$S = C \log L \quad (1)$$

(S は心理量, L は標準刺激量, C は定数)

よって刺激量の対数に比例させて心理量を変化させる。

また本稿では、単発的な攻撃の成否や試合の勝敗のみではなく、過去一定時間以内の攻撃成否回数や過去数試合における勝敗数を逐次記録していき、快-不快パラメータの変化起因に用いる。

i 攻撃(被攻撃)の成功とその威力

攻撃を当てるとうれしい、当てられると悲しい。それは攻撃の威力が大きいほど、また残りライフが少ないほど、大きな快-不快の心理状態変化を引き起こすと考えられる。読み取ったダメージの大きさに比例させて快-不快を正負に変化させる。また、特殊な映像的演出がなされる大技や、一連の攻撃終了後に「Great!」などの文字情報が画面に出現する攻撃に関しては、ダメージの大きいとともに視覚的にも通常より強く訴えられるため、心理的に大きな影響があると考ええる。 $Life_{old}$ を数フレーム前の残りライフ率(%), $Life_{now}$ を現在の残りライフ率(%), c を大技や文字情報が出現した場合の定数, a_1 を係数として、以下の式を設定する。

$$m_1 = a_1 (\log Life_{old} - \log Life_{now}) + c \quad (2)$$

($Life_{old}, Life_{now} > 0$ 攻撃時 $a_1 \geq 0$,
被攻撃時 $a_1 < 0$)

ii 一定時間内の攻撃(被攻撃)回数

何度も攻撃が連続で成功する(受ける)場合は、ある程度交互に攻撃被攻撃を繰り返す場合と比較して、大きなうれしさや悲しさを得、快-不快が変化する。 n_1 を一定時間前から現在までの期間の攻撃成功(攻撃受け)回数, a_2 を係数として、以下の式を設定する。

$$m_2(t) = a_2 \log n_1(t) \quad (3)$$

($n_1 > 0$ 攻撃時 $a_2 \geq 0$, 被攻撃時 $a_2 < 0$)

iii 予期した攻撃をはずす(回避する)

特殊な映像効果で演出される大技は、通常攻撃と比較して非常に強力な威力を持つように設定されているため、使える回数が制限されている場合がほとんどであり、3.1 節②の反応発話2が発生するような状況では、攻撃側は残念に思い、被攻撃側は喜ぶ。 a_3 を係数として、快-不快パラメータ変化の一要因として随時判定する。

$$m_3 = a_3 \quad (4)$$

(攻撃時 $a_3 < 0$, 被攻撃時 $a_3 \geq 0$)

iv 残りライフゲージが大きく開いている

両者の残りライフゲージを比較し、大きく差が開いている場合、勝っている方は上機嫌、負けている方は不機嫌になる起因となる。 $Life_{player}$ をプレイヤーの残りライフ率(%), $Life_{adversary}$ 残を仮想対戦プレイヤーの残りライフ率(%), a_4 を係数とし、以下の式を設定する。

$$m_4(t) = a_4 (\log Life_{player}(t) - \log Life_{adversary}(t)) \quad (5)$$

$$(Life_{player}, Life_{adversary} > 0)$$

v 試合に勝ち続ける、または負け続ける

試合に勝つとうれしい、負けるとくやしい。もともと単純な快-不快変化の起因である。 $N_{adversary}$ を仮想対戦プレイヤーの勝ち数、 N_{player} をプレイヤーの勝ち数、 a_5 を係数として、以下の式を設定する。

$$C_1 = a_5 (\log N_{adversary} - \log N_{player}) \quad (6)$$

$$(N_{adversary}, N_{player} > 0)$$

i ~ v により、快-不快パラメータ値 M は以下のよう定まる。

$$M(t) = m_1 + m_2(t) + m_3 + m_4(t) + C_1 \quad (7)$$

$$(-100 \leq M(t) \leq 100)$$

ただし C_1 は1試合に1度試合開始直前に定数として加算、 m_1, m_3 は随時、それ以外は試合中に単位時間ごとに判定する。

4. 2. 快-不快パラメータと快・不快ランクの対応付け

快-不快ランクと発話音声ファイルの対応付けに関しては3章のとおりだが、ここでは快-不快パラメータと快・不快ランクの対応付けについて説明する。特定の快-不快パラメータ値を閾値として快-不快ランクと対応付けする方法では、プレイヤーに発話音声ファイルのランク切り替わりがはっきり分かってしまい、人間らしい境界のあいまいさに欠けると考える。そこでパラメータ値とランクの対応付けは図3のようなメンバーシップ関数を設定し、グレード値をそのランクの発話音声ファイルの選択確率とする。これにより快-不快ランクの境界の変動性を持たせる。

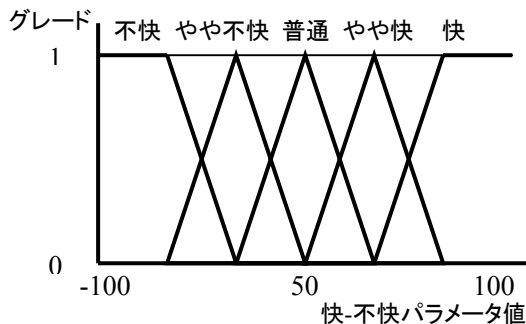


図3 パラメータ値とランクの対応

4. 3. 発話頻度パラメータ値の算出

ゲームプレイ中の発話頻度は、ゲームへの集中の度合いや没入の度合い、また興奮に依存して上下する。それらを定量化するための発話頻度パラメータを設定し、仮想対戦プレイヤーの音声ファイル出力確率と対応付け、発話頻度パラメータ値が高いほど発話頻度が増す。ビデオより発話頻度パラメータを変化させる要因として本稿では、残りライフ量、両者の残りライフの差、残り時間、試合の勝敗数を考え、以下の式により発話頻度パラメータ値を算出する。

I 残りライフ量

試合開始当初、残りライフが多いときは、緊張感がさほどないが、残りライフが少なくなると、緊張が高まり没入感が高くなる。それによって発話頻度が増える。 $AveLife$ を両者のライフ残量率の平均値(%), b_1 を係数として、以下の式を設定する。

$$f_1(t) = b_1 (\log 50 - \log AveLife(t)) \quad (8)$$

II 両者の残りライフの差

実力伯仲で、両者の残りライフの値が近いときには、プレイヤーは興奮し、発話頻度が高くなる。逆に両者の間に力の差が大きくあり、残りライフの差が大きいときは、興奮めしてしまい飽きが早くなる。 b_2 を係数として、以下の式を設定する。

$$f_2(t) = b_2 (\log 50 - |\log Life_{player}(t) - \log Life_{adversary}(t)|) \quad (9)$$

$$(Life_{player}, Life_{adversary} > 0)$$

III 残り時間

試合終了間際になって残り時間が少なくなると、緊張が高まり発話発生が多くなる。 $Time$ を残り時間、 $Time_{MAX}$ を残り時間の最大値、 b_3 を係数として、以下の式を設定する。

$$f_3(t) = b_3 (\log Time_{MAX} - \log Time(t)) \quad (10)$$

$$(Time > 0)$$

IV 試合に勝ち続ける、負け続ける

両者の間の実力の差が大きく、片方が一方的に勝ち続けると、つまらなく感じ、ゲームの飽きが早い。それに伴い発話頻度も減少する。 b_4 を係数、 α を定数として、以下の式を設定する。

$$C_2 = b_4 (-|\log N_{adversary} - \log N_{player}| + \alpha) \quad (11)$$

$$(N_{adversary}, N_{player} > 0)$$

I ~ IVにより、発話頻度パラメータ値 F は以下のよう定まる。

$$F(t) = f_1(t) + f_2(t) + f_3(t) + C_2 \quad (12)$$

$$(0 \leq F(t) \leq 100)$$

ただし C_2 のみ1試合ごとに1回、試合開始直前に定数

として加算し、それ以外は試合中に単位時間ごとに判定する。

5. ゲーム状況の読み取り

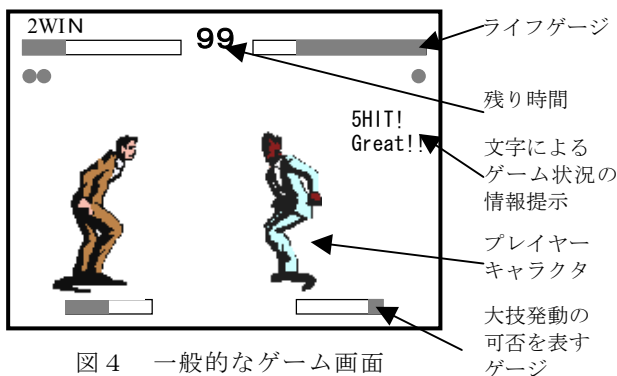
ゲームの状況読み取りには、実際に市販のテレビと家庭用ゲーム機で市販のゲームをプレイしている映像を、AVケーブルとキャプチャボードによって実時間でPCに取り込み、その映像から特定のゲーム状況を表す特定の情報を検出する方法を用いる。

この方法のメリットとして、ゲーム開発に携わりゲームの内部プログラム情報を参照することができなくても、数多くのゲームに応用できる点が挙げられる。他に、コントローラの入力情報を読み取るような特殊なハードウェアを用いることなく実行できるというメリットもある。

格闘アクションゲームの対戦画面は、図4のような伝統的なスタイルが存在する。ゲームそれぞれによってのアレンジはあれど、これに従うものが多い。また対戦中には、一撃で大ダメージを与える大技の発生前や発生中に、画面反転や時間停止などのなんらかの特殊映像効果が用いられ、連続攻撃やカウンター成功などによってプレイヤーの操作技術が高いと評価した場合に、それを画面中に文字を出すことによって伝えたりすることが一般的である。

このように格闘アクションゲームには、ゲームの状況を視覚的にフィードバックする一般的な映像上の性質が存在する。これは最新のゲームになるほど情報量が多くなる傾向にある。これにより比較的容易に映像から様々なゲーム状況を読み取ることができる。

具体的に本稿では、ゲーム映像中の特定箇所を指定し、その箇所の画素値の変化からゲーム状況を読み取る。



5. 1. 映像からのゲーム状況読み取り

ゲーム映像中の図5で説明する位置に1ドットのマーカをつけて、画素値を読み取り、これがあらかじめ設定した特定の値になった場合にフラグを立て、このフラグの組み合わせによりゲーム状況を判断する。格闘アクションゲームのゲーム状況は、数分の1秒単位で大きく変化していく

ため、プログラム実行速度をできるだけ早くする必要はある。それゆえ、処理速度の遅い画像のマッチング等ではなく、このマーカをつける方法を用いる。またゲーム機から取り込んだ映像は、カメラ撮影によるものと比較してノイズが少なく、広範囲の読み取りをする必要がないのも理由である。

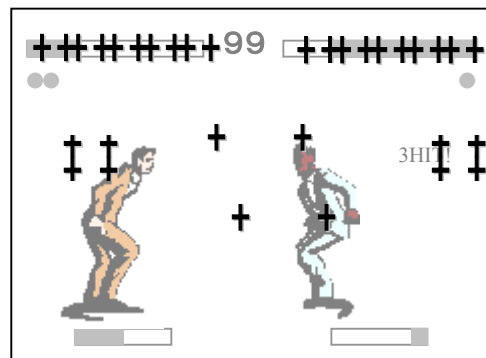
具体的には、まずプレイヤーと仮想対戦プレイヤー両方のライフゲージをそれぞれ20等分するようにマーカを置く。あらかじめ設定しておいた残りライフ部の画素値を読み取った場合は1、それ以外の場合は0と出力するようにし、これの20個ずつの組み合わせによって5%単位で残りライフを判断する。また1フレーム前との差分により、瞬間ダメージ量が判断できる。

試合の開始と終了は、どちらかのプレイヤーのライフゲージに置いたマーカ全てが0を出力、かつ中央部のマーカで試合終了を表すエフェクトを検出した場合に判断できる。

今回扱った格闘アクションゲームでは、大技の発生前の瞬間に、一部を除き画面全体の明度彩度を非常に低くする演出がなされる。これより、マーカの大部分が低い画素値を検出した場合、大技発生の直前と判断する。その直後に被攻撃側がダメージを受けた場合、大技の成功の判断ができる。または大技成功のあとにはほとんどの場合に画面中部両端の文字情報提示ゾーンに連続攻撃回数やGreat!などの文字情報が表示される。これをマーカの画素値変化により読み取ることも大技成功の判断が可能である。

3.1節の④文章発話が発生しやすい中休み状態の読み取りは、まず予期、攻撃、被攻撃中ではない状態をそれとみなす。また、今回の格闘アクションゲーム中の、大技成功の直後は被攻撃側はキャラクターが倒れてしまって、コマンド入力できないダウン状態になる性質を利用して判断する。互いに一定時間ダメージがない場合を小休止状態とみなすこともできる。

残り時間については、残り時間の最大値と、試合開始から試合終了までのプログラム実行回数を対応付けすることによって判断する。



5. 2. ゲーム状況による分類との対応付け

このようにして読み取ったゲーム状況と、3.1 節で分類した発話音声ファイルに対応付ける。

① 予期発話との対応付け

前節で説明したように、大技の発生直前には、画面の明度彩度が非常に低くなる。複数のマーカーよりその演出効果を読み取った場合に予期発話音声ファイルと呼び出す

② 反応発話との対応付け

仮想対戦プレイヤー側のライフゲージ部分の画素値を、複数のマーカーから読み取り、残りライフ部分に2フレーム前、1フレーム前との差分があればダメージを受けたと判断し、反応発話音声ファイルと呼び出す。現在のマーカーの画素値と過去2フレーム分の画素値の差分を取ることによって、キャラクターなどによって1フレーム間ライフゲージが隠れてしまったり、ノイズによって1フレーム間不正な画素値を検出してしまった場合に対応できる。

また予期した攻撃が実際に成功したか失敗したかの判断は、①の予期発話呼出し後、数フレーム以内にダメージがあるかないかによって判断する。

③ 能動発話との対応付け

プレイヤー側のライフゲージ部分の画素値を、複数のマーカーから読み取り、残りライフ部分に2フレーム前、1フレーム前との差分があればダメージを与えたと判断し、反応発話音声ファイルと呼び出す。

人間は能動発話を厳密には攻撃を行う一瞬前に発生する機会が多いが、この方法により数フレーム発生が遅れた場合も、遅れた感覚は感じなかった。むしろ次章の評価実験で、被験者が仮想対戦プレイヤー側のキャラクターを操作した結果、自分が言おうとしたことを、一瞬先に仮想対戦プレイヤーに言われてしまうといった感覚を得ており、処理遅れについては全く問題ないと考える。

④ 文章発話との対応付け

文章発話の中には非常に特定された状況、例えば試合終了の瞬間に限定された発話内容やライフ差が大きく開いているときに限定された発話内容など、がある。これらはファイル1個単位でゲーム状況との対応付けを行う。試合終了の瞬間なら、中央部分のマーカーが試合終了時の特殊エフェクトである「K.O.」を読み取った瞬間と対応付ける。

それ以外の文章発話、ある程度状況によらず発生するものについては、決まった対応付けはせず、上記①②③の発話が発生する状況以外の中休み状態などのタイミングで、ランダムで発生させる。

5. 3. 読み取り正確さの評価

実際に数十分、仮想対戦プレイヤーが実装された格闘アクションゲームをプレイし、読み取りの正確さの評価を行った。なお、テレビ上のゲーム画面と、PCから出力される残り生命などの情報のタイムラグは、0.06秒前後であり、体感では差を感じられなかった。

読み取る情報として最も重要な残りライフは、プレイヤーキャラクターなどによってライフゲージの残りライフと蓄積ダメージ量の境界部分が数フレーム以上隠れてしまう場合を除き、ほぼ100%の精度で読み取ることができた。この読み取れない場合だが、評価中に数回しか出現せず、無視してもよいと考える。

大技の発生については100%の精度で検出できたが、成否に関しては、技の発生からダメージ発生まで時間がかかる場合、また攻撃の一部しか相手に当たらなかった場合に誤検出が発生した。

文字情報の読み取りに関しては、プレイヤーキャラクターや必殺技の映像や背景映像の位置変化によっては、文字情報の画素値と混同してしまう場合があり、総合すると7割程度の精度になった。

これらの結果より、検出ミスや誤検出、または不適切な対応付けによって、適切ではない発話音声ファイルが2割程度出力された。ほとんどの場合、発話内容の解釈の仕方を変えることで受け入れられる範囲であったが、中には大きな違和感を感じるものもあった。より細かなゲーム状況の読み取りと、発話音声ファイルの分類方法考案が今後の課題である。

6. 心理評価

当システムの実装された格闘アクションゲームを大学生男女15名の被験者が体験し、アンケートに答えた。アンケートの質問内容は以下のとおりである。

- ① COMが発話するのは楽しい
- ② 発話がないときと比較して、ゲームを面白く感じる
- ③ 発話がないときに比べて、長くプレイできたと思う
- ④ COMの心理状態変化が感じられた
- ⑤ またやってみたいと思う
- ⑥ 違うゲームに実装されたらまたやってみたい
- ⑦ 違う人物の声が実装されたらまたやってみたい

アンケートの回答項目は以下のとおりである。

- 1 とてもそうは思えない
- 2 そうは思えない
- 3 なんともしえない
- 4 そう思う
- 5 とてもそう思う

結果を図6に示す。

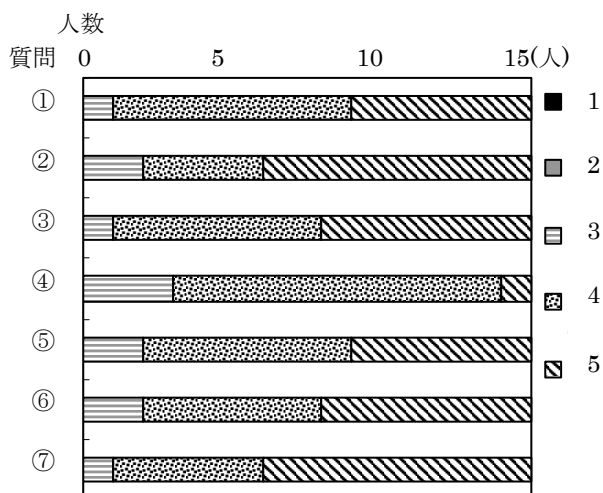


図6 アンケート結果

また自由記入形式の質問では、主に次のような感想が得られた。

- ・ 状況判断がよい。
- ・ タイミングがよい(反応が早くストレスを感じない)。
- ・ 声の抽出もとの人物と対戦しているような錯覚を得た。
- ・ COM が喜ぶ声にむかっついて、次こそは勝とうという気持ちが生まれる。
- ・ 感情が設定されていて、変化することによって音声も変わるのがおもしろい。
- ・ COM が感情的発話をするのにはつい笑ってしまう。
- ・ 明らかに不適切な声を出したときは、不快感が湧く。
- ・ 知人の声実装されたものをやってみたい。

アンケート結果より、仮想対戦プレイヤーはプレイヤー1人で対戦型ゲームを行う際の面白さの改善において有意であることが確認できた。

特筆すべき点として、①から⑤の質問に対しては、音声ファイル作成元の人物の知人友人である被験者のほうが全く面識のない被験者に比べて高い評価を残した。また質問⑦に対しては、面識のない被験者のほうが高い評価を残した。この結果から、発話音声ファイル抽出元のプレイヤーと知人友人関係であるプレイヤーのほうが、当システムをおもしろく感じる事が分かる。

これにより、さらにプレイヤー全体からの評価を上げる手段としては、声の抽出元のプレイヤーを、プレイヤーの大半が知っているであろう芸能人や有名人に変更することが考えられる。

7. まとめと今後の課題

本研究では、ゲーム映像からゲーム状況を判断し快-不快と没入度の内部パラメータを変化させ、ゲーム状況とパ

ラメータに対応付けした発話音声ファイルを出力することによって、人間らしい感情的発話をする仮想対戦プレイヤーを生成し、市販の格闘アクションゲームに実装した。

心理評価の結果、当手法は対戦型ゲームをプレイヤー1人でプレイする場合に、面白さの改善において有意であったといえる。

今後の課題としては、現在は読み取れるゲーム状況の種類が少ないので、音声認識など別の検出方法を取り入れることが考えられる。また、発話音声ファイルの作成および分類を現在手作業で行っており、手間がかかっているため、一部自動化をしたい。そして今回は格闘アクションゲームのみを扱ったが、これ以外の対戦型ゲームに仮想対戦プレイヤーを一般化したい。

他に、よりシステムの効果を上げる為に、仮想対戦プレイヤーの声の抽出元にどのような人物を用いるか、どのような発話音声ファイルの内容やタイミングによってプレイヤーが面白いと感じるか、といった心理的な検証も課題である。

参考文献

- [1] AIBO Official Site, <http://www.jp.aibo.com/>
- [2] 株式会社エイ・ジー・アイ, <http://www.agi-web.co.jp/>
- [3] 田中彰人, 星野准一, “人間の模倣によるアクションゲーム AI”, 情報処理学会, エンタテインメントコンピューティング 2005, pp.76-81, 2005.
- [4] 黒宮寧, 磯田佳徳, 長沼武史, 倉掛正治, NTT ドコモマルチメディア研究所, “日常的なインタラクションによるユーザーエージェントの実現”, 第 17 回人工知能学会全国大会, 2003.
- [5] Hiroki Ogawa, Tomio Watanabe, “InterRobot: speech-driven embodied interaction robot”, Advanced Robotics, Vol.15, No.3, pp.371-377, 2001.
- [6] 深山 篤, VincentBao Pham, 大野 健彦: 視線分析に基づく擬人化エージェントのユーザビリティ評価の検討, 電子情報通信学会技術報告, HIP2003-136, 2004.
- [7] サリヴァン. H. S. 著, 中井 久夫 翻訳, “精神医学的面接”, みすず書房, 1986.
- [8] 土屋昭二, 竹村和久 編, “感情と行動・認知・生理感情の社会心理学”対人行動研究シリーズ④, 誠信書房, 1996.
- [9] 村岡哲也, “心理物理学-心理現象と視機能の応用-”, 技報堂出版, 2005.
- [10] 御領謙, “人間の情報処理”, サイエンス社, 1985.