

感情表現生成を行う仮想ゲームプレイヤー

Virtual Adversary Player by Generating Emotional Utterance

白鳥 和人^{†‡} 塩入 健太[†] 星野 准一[†]

Kazuhito Shiratori^{†‡}, Kenta Shioiri[†], Junichi Hoshino[†]

[†]筑波大学システム情報工学研究科 [‡]科学技術振興機構

[†]University of Tsukuba Graduate School of System and Information Engineering

[‡]Japan Science and Technology Agency

概要: 本稿では, 格闘対戦型ゲームの自動対戦キャラクターの仮想プレイヤーにゲーム状況やプレイヤー行動に応じた発話音声を加える手法を提案する. 本提案ではまずゲーム状況を画面内マーカー位置の画素変化からライフポイントや技パターンを読み取り獲得する. これに基づき計算機側プレイヤーの感情パラメータを算出する. 音声会話データは実際にゲームプレイ中の発話を録音し抽出しパラメータと対応付けておく. このシステムで遊ぶときこれらの会話セットを適切に選択し仮想プレイヤーに発話させる. 実際のユーザにプレイをさせた結果, 対戦型ゲームを計算機プレイヤーと行う場合での手法の有用性が確かめられた.

Abstract: This paper proposes a virtual fighting game player who speaks on various situations from game progress. The method of capturing a game status of Life-point-gauge and skill-taking-patterns is read from special pixels on game video graphic frames. Emotional parameters are calculated from this value. Utterance sound is captured and classified as vice-set at on same time. After those preparations, Game is played with this system. Appropriate voices made from real speech on playing person is selected voice-set by situation. A user evaluation shows this method adds game joy increased.

キーワード: エンタテインメント, ゲーム演出, 自動対戦, 発話生成, プレイヤー感情分析

Keywords: game direction, automatic player, Utterance generation, game player emotion analyze

1. はじめに

エンタテインメントは私たちの生活を豊かにするために重要な役割を果たしている. コンピュータを用い現実のスポーツをシミュレートし楽しむという形式は, 実参加することが困難な高度なスポーツの楽しさを味わうことが出来るという効果がある. こういったコンピュータエンタテインメントの中でもディスプレイ画面にキャラクターを登場させ対戦させる拳闘・柔剣道・スポーツ喧嘩などの格闘ゲームはや家庭用ゲーム機向けなど, 多くのタイトルが制作されており, その人気は衰えていない.

対戦ゲームの特徴としてゲーム画面内のキャラクターに指示を与え, 技を繰り出して攻撃をしたり, それを回避させたりする対峙の構造がある. 対戦する場合, 通常は相手プレイヤーが人間の[人対人]であるが, 遊び方のバリエーションとして, プレイヤーは通称COMと呼ばれるゲームシステムが提供するAIと対戦する

ことも出来る. ところが[対COM戦]のプレイモードに比べつまらなさと感じるプレイヤーが少なからずいる. これらのタイトルではプレイヤーのスキルアップやプレイ方法の一形態として重要であるにも関わらず, もっと遊んでもらうようにする方向に着目した仕組みは十分に用意されていない.

本稿ではこの点を改善する一案として格闘の対戦COMをプレイしているとみなす仮想の人物を用意し面白さの向上に寄与できる発話付加の枠組みを提案する. これは[対COM戦]のCOMを操作しているプレイヤーを模擬するもので, ゲームプレイ状況を分析して優劣状況に応じた感情評価に基づく発話をするものである. これにより, 人間のプレイヤーが一人で遊ぶ場合の楽しさを広げる可能性が加えられる.

1.1. 従来研究

対戦ゲームの面白さ向上策としてはキャラクターの種類や技, 効果音などの演出を加えるほか背景世界

観の深い設定などがよく行われる。登場キャラクターへの演出だけでなくコンパイル社(株)のパズルゲーム「ぷよぷよ」のように、有利・不利の状況を判断して簡単な発話による反応を組み込む方法も考案されている。しかしこれらのゲームキャラクターの反応は単調であり人間と対戦しているような印象を得るのは難しい。近年、人間からの干渉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いる。Breazeal et al.[1]は、人間の感情モデルを用いて、表情豊かな顔表情ロボットを作り出した。Takanishi et al.[2]も、ヒューマノイドロボットを用いた感情表現についての研究を行っている。さらに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[3][4]がある。しかしこれらに対戦ゲームに適用するのは難しい。ほかに、対戦ゲームのAIに関する研究では、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[5]、対戦相手からの感情的な反応の発話については触れていない。

以上のように COM を相手にした事例ではギャラリー 応援やゲーム内キャラクターの人間表現に関心が向けられ、ゲームの相手をしてくれる存在の提案はされていない。

2. 仮想対戦プレイヤーの概念

2.1. 仮想対戦プレイヤーの定義と要件

仮想対戦プレイヤーとはゲーム中の相手キャラクターではなく、そのキャラクターを操作していると想定される相手即ち COM に対して擬似的に人格を付与したものを指す。通常 COM は人格が無くひたすら無機的に対戦相手キャラクターを操りゲームの相手を務める。もしこの存在が感情をもち相手を務めてくれたならプレイヤーが一人でゲームをする場合の楽しさが増すのではないかと考え本稿のシステムを提案する。

図1がそのイメージである。

[人対人] プレイでは対戦中状況に応じた発話が

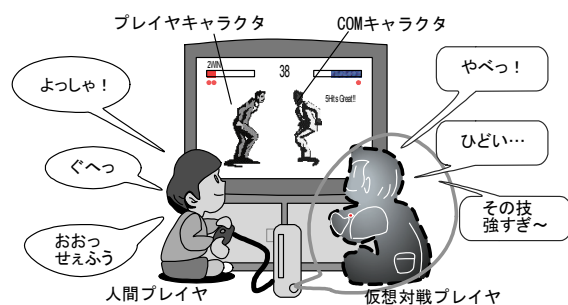


図1 仮想対戦プレイヤーの概念

交わされるが計算機相手では登場キャラクターの音声や効果音こそ出るものの仮想の相手である[COM]は沈黙している。実際の間人が対戦相手であれば発話によって対戦中に様々な反応や感情が表出する。この感情はプレイ中のゲーム進行状況は似通っていても、以前の戦況での勝敗連鎖状態などそこに至るまでの経緯によって異なる発話を行うのが常である。実際にプレイの様子を観察してみると例えば“ゲームに負けた瞬間”の場合連勝のあとたまたま負けたのであれば「あれ、まあいっか」のような発話をするが、連敗後、重ねて負けると「もういやだ、やめたい」というような発話をする様子が見られる。多様な反応と感情表出をする発話機能搭載は対戦型ゲームのエンタテインメント性の重要な要素である。

これらを念頭に置くと、仮想のプレイヤーは単に戦いのメッセージを放つだけではなく、ゲーム状況を判断してそれによって生ずる心理変化を再現したヒューマンライクな応答が必要であることがわかる。応答についてはゲームプレイ中であることを考慮すると音声による発話が入への情報伝達の同時性という点で好ましい。

2.2. システム構成

システムは以下のサブシステムからなる。

- 1) ゲーム状況を読み取る機能
- 2) 仮想プレイヤーの感情モデル
- 3) モデルから算出した発話頻度モデル
- 4) 一定のルールで言葉を選定する発話選択機構
- 5) 分類済みの発話音声データ群

図2がその大まかな構成である。ゲームプレイ時には発話音声データごとの発話頻度の初期値と、発話頻度パラメータを参照して、確率に応じた発話選択ルールに基づき音声出力を行う。音声はゲームプレイ時に用いる分析方式とほぼ同じ仕組みであらかじめ収録・分類を行い保管しておく。

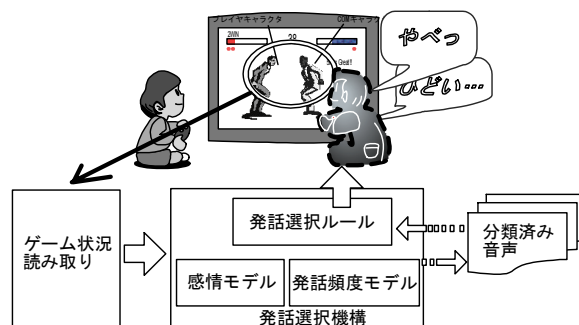


図2 システム構成

2.3. ゲーム状況の分析と分類

ゲーム状況とは対戦キャラクターの体力など増減や戦況の進退などのように対戦中の双方の優劣状態を指す。これらは一定の値に達したら有利か不利かというふうに一元的には定まらず、大技を決められて一挙に体力が減った場合や、制限時間終了間際か否か等の要件も含まれる。

しかしこれらに対戦型ビデオゲームという環境に適応させることは難しい。たとえばそのためにゲーム機内のソフトウェア内部の変数値を提示するのが単純だが、ゲームプレイのチートプレイ(ずるいプレイ)に利用されることを開発者が嫌い、公式に提供してもらうことは現状では難しい。また、プレイヤーに公平感を抱かせる意味からも人が判断するのと似た条件であることが好ましい。したがって基本的には映像や音声などゲーム機から正式に出力されている信号を捕捉して利用するのが良い。知るべきゲーム状況は「変化」要素が最も多いので映像情報からの取り込みが最も適している。

映像で得られ、かつプレイヤーの心理に影響を与えらると思われるゲーム状況はおおよそ下記の事柄である。但しここでは最も原始的な1対1の対戦で攻撃によりダメージを与えて規定のダメージ耐量、つまりライフポイントをゼロにした側が勝つという形式の対戦ゲームを前提にする。

- 1)双方のキャラクターのライフゲージの量
- 2)同様にその単位時間当たりの変化量
- 3)残り時間
- 4)大技などが決まったタイミング

ゲーム状況はプレイヤーの理解を容易にするため図3のように一定の位置に表示することが多い。これを心理パラメータとして利用する場合なら十分な速度で取り出せる。また一撃で大ダメージを与える大

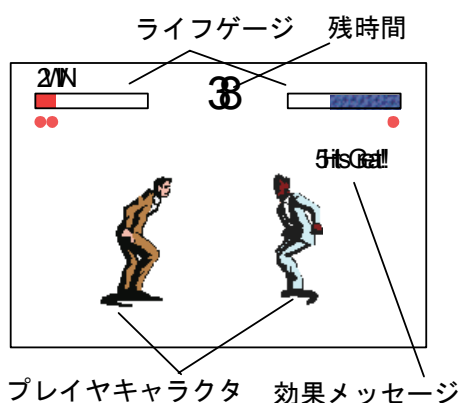


図3 典型的なゲーム画面

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たり、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ため4)のような事項も抽出できる可能性がある。

以上から比較的容易に映像から様々なゲーム内状況を読み取ることができる。

3. 感情モデルと発話生成手法

3.1. プレイ中の発話の分類

人がプレイ中にする発話と同じものを発話させるのであるから、どのような発話がなされるのかを明らかにする必要がある。感情的発話は発生時のゲーム内状況に応じて大別すると4属性に分類できる。

a) 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、単語単体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。この発話の発話頻度は最初の時点で高い。

例) やば、うわっ、危ない、よっしゃ、いける

b) 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である。攻撃側と被攻撃側それぞれ予期の有無と攻撃の成功失敗によって3つに分類する。表1に被攻撃の場合を示す。予期のある場合は発話頻度の初期値が高い、ない場合は低い。

c) 能動発話

攻撃をする際の掛け声。命令調の動詞や技名などの固有名詞であることが多い。ゲーム内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。発話頻度の初期値は低い。

例) くらえ、〇〇拳

d) 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。

試合終了時にライフに大きく差をつけられている状況のような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。前者は分類ごととくくりゲーム内状況に対応付けることは困難であるので、発話音声データ個別にゲーム内状況に対応付ける。発話頻度の初期値はデータごとに異なる。

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、やってみよう、ミスった

e)その他

a)~d)の分類外でゲームプレイに関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?

a),b),c)の発話音声データはその分類単位で、上で説明した出現の傾向に合わせてゲーム場面に就対応付ける。d)の文章発話のうち、特定の状況に起因するのは、データ1個単位でゲーム場面と対応付ける。状況に関わらず a),b),c)以外のケースで発せられる言葉はランダムに呼び出す。

3.2. 仮想プレイヤーの感情モデル

前節の発話はプレイヤーの心理状態に応じて選択させるために、仮想プレイヤーに感情モデルを設定しここから選択条件を決定する。

ゲームプレイ中、プレイヤーの感情は動的に変化する。これを知るためにゲームプレイの様子を撮影しどのようなものであるか考察した。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す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]。例えば語尾に撥音をつける、語尾を伸ばす、低くゆっくり発話する、高く早く発話するのでは聞き手の聴取印象はかなり異なる。

本稿では発話音声データを、聴取印象をもとに“快-不快”の尺度で分類する。感情のモデルにはプルチックのモデル[8]など様々なものがあるが、本稿の対戦型ゲームプレイ中に限るなら、キャラクターの行動

の自由度が高いRPG等のゲームジャンルと違い、操作しているキャラクターの行動が「格闘」に制限されているので多種多様で複雑な心理は生まれにくいと仮定し、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分類である快-不快とした。なお全ての感情は、快-不快の何れかに分類できる。

具体的適用では不快、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ると定義する。発話音声データの快-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表2に示す。

表2 ゲーム内状況分類と快-不快分類

	予期発話	反応発話	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	えいっ	うおお
快	お	うまいっ	やー	うまいね

3.3. 心理パラメータの定性的考察

心理パラメータは快-不快パラメータを要素に持つ。また感情表現としては何を話すかということの他、たくさん話すかどうかも重要である。そこで発話頻度パラメータも感情モデルの要素とする。

たとえば連続して勝ち続けており非常に快の状態でも、対戦相手に対する同情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、イラっとして暴言が多くなるといった状況が該当する。

本稿では仮想対戦プレイヤーの心理状態を変域±50の符号つき整数の快-不快パラメータで表し、これを快から不快までランク分しておいた発話音声データと対応付ける。発話頻度パラメータは発話頻度に正比例する変域0から100の整数値とした。各パラメータの変動範囲と分解能は今回扱う格闘アクションゲームを実際に数人が「人対人」でプレイした予備実験の意見聴取で決定した。

ゲームから取得した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との関係は、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

表1 反応発話の分類(被攻撃側の例)

	攻撃を受けた	回避した
予期あり	反応発話1	反応発話2
例	「うえ」, 「まじでえ」	「セーフ」 「助かった」
予期なし	反応発話3	該当なし
例	「うお」,「む」,「お」	

つというウェーバー・フェヒナーの法則[9]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)

つまり刺激量の対数に比例させて心理量を変化させるのが感覚に適合する。

3. 4. 快-不快パラメータ

心理状態の快-不快可変要因として, 成功又は受けた攻撃の威力, 一定時間内の攻撃又は被攻撃の回数, 予期した攻撃を外す又は回避する, 残りのライフゲージが大きく開いている, 試合に勝ち又は負け続ける, 等を考慮する. 具体的には $i \sim v$ に示す式によって快-不快パラメータ値逐次求める. 更に単発的な試合の勝敗のみではなく, 過去数試合における勝敗数を累積記録していき, 快-不快パラメータの可変要因とする.

i 攻撃(被攻撃)の成功とその威力

攻撃を当てるとうれしく, 当てられると悲しい. それは攻撃の威力が大きいほど, またそのときの残りライフポイントが少ないほど, 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. n_1 を攻撃(被攻撃)の回数, $Life_{old}$ を数フレーム前の残りライフポイント率(%), $Life_{now}$ を現在の残りライフポイント率(%), 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に基づき変化を求める.

$$m_1(n_1) = m_1(n_1 - 1) + a_1(\log Life_{old} - \log Life_{now}) + c \quad (2)$$

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

ii 予期した攻撃を外す(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, 3.2 のb) 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. n_1 を予期した攻撃を外した(回避した)回数, 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.

$$m_2(n_2) = m_2(n_2 - 1) + a_3 \quad (3)$$

(攻撃時 $a_3 < 0$, 被攻撃時 $a_3 \geq 0$)

iii 残りライフポイントが大きく開いている

両者の残りライフポイントゲージを比較し, 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りやすい. t を時間, $Life_{player}$ をプレイヤーの残りライフポイント率(%), $Life_{now}$ を仮想対戦プレイヤーの残りライフポイント率(%), a_4 を係数とし, 以下の式に従い変化を与える.

$$m_3(t) = m_3(t-1) + a_4(\log Life_{player}(t) - \log Life_{now}(t)) \quad (4)$$

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

iv 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, 負けるとくやしいといったもっとも単純な快-不快変化である. $N_{adversary}$ 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, 以下の式で変化させる.

$$C_1 = a_5(\log N_{adversary} - \log N_{player}) \quad (5)$$

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

$i \sim iv$ を纏めると, 快-不快パラメータ値 M は以下のように定まる.

$$M(t) = m_1(n_1) + m_2(n_2) + m_3(t) + C_1 \quad (6)$$

($-50 \leq M(t) \leq 50$)

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_2 は随時, m_3 は試合中に単位時間ごとに判定する.

3. 5. 快-不快ランク

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データ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. そこでパラメータ値とランクの対応付けは図4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データの選択確率とする. こうして快-不快ランクの境界に揺らぎを設ける.

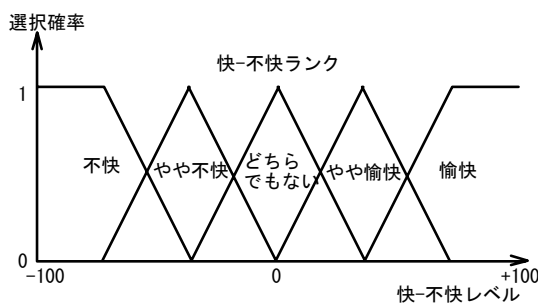


図4 パラメータ値とランクの対応

3.6. 発話頻度パラメータ

ゲームプレイ中の発話頻度は、ゲームへの集中や没入の度合い、興奮度に応じて増減する。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データ出力確率と対応付ける。発話頻度パラメータ値と発話頻度は正比例とする。最終的に仮想対戦プレイヤーに発話させるか、させないかはこの発話頻度パラメータと分類ごとの発話頻度の初期値を参照して決定する。

撮影したビデオの観察から発話頻度パラメータの可変要素には、残りライフポイント量、両者の残りライフポイントの差、残り時間、試合の勝敗数を加味するのがよいと判った。そこで以下のルールによって発話頻度パラメータ値を算出する。

I) 残りライフポイント量

試合開始当初、残りライフポイントが多いときは緊張感がさほどないが、残りライフポイントが少なくなると緊張し感情移入する、その結果発話頻度が増える。AveLife を両者のライフ残量率の平均値(%), b_1 を係数として頻度パラメータを次式で変化させる。

$$f_1(t) = f_1(t-1) + b_1(\log 50 - \log AveLife(t)) \quad (7)$$

II) 両者の残りライフポイントの差

実力仲間で両者の残りライフポイントの値が近いとプレイヤーは熱狂し発話頻度が上がる。逆に両者の実力に開きがあり、残りライフポイント差が大きいと興ざめし飽きが早くなる。 b_2 を係数として、次の式を設定する。また両者の残りライフポイントが近い、開いているかの判断は、片方のプレイヤーの残りライフポイントが 60%、もう片方のプレイヤーの残りライフポイントが 40%のときの心理量を閾値として、差がこれ以下であれば両者の残りライフポイントの差は近くこれ以上であれば開いているとみなす。

$$f_2(t) = f_2(t-1) + b_2((\log 60 - \log 40) - |\log Life_{player}(t) - \log Life_{now}(t)|) \quad (9)$$

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

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、TimeMAX を残り時間の最大値、 b_3 を係数として、次式で表される。

$$f_3(t) = f_3(t-1) + b_3(\log Time_{MAX} - \log Time(t)) \quad (10)$$

(Time > 0)

IV) 試合に勝ち続ける、負け続ける

両者の実力差が大きく一方的な勝負であるとならなく感じ飽きが早く来る。そうなると発話頻度も減少する。 b_4 を係数、 α を定数とし次式で表される。

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

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

これらから発話頻度パラメータ値 F の単位時間当たりの次式の纏められる。

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

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

ここで C_2 のみ 1 試合に 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。

なお、この発話頻度パラメータと、前述の快-不快パラメータが実際にどのような動き方をするかについては、次章でその例を示す。

4. システム実装例

前章で述べたシステムの有用性を確認するため以下の手順と方式で仮想格闘対戦プレイヤーを実装した。

4.1. 発話データの作成

まず発話音声の収録と分類を行う。大まかな手順は以下の通りである。

【ゲームプレイ状況の収録】

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撮影し、同時に発話音声の録音をする。録音した音声を切り出し発話音声データファイルセットとする。

【発聲音声の分類】

ビデオを参照し発話音声データを発話発生時のゲーム内状況によって 4 分類、さらに音声から感じる快-不快の印象によって 5 分類する。この分類ルールは 3.2 節の表 2 で示したものである。ゲーム内状況による分類ごとに発話頻度の初期値を設定する。

【心理状態変化のマーキング】

仮想対戦プレイヤーの内部パラメータとして心理状態を表す快-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。さらにゲーム内状況とパラメータ値変化を対応付ける。

【分類音声と心理パラメータ対応付け】

快-不快パラメータと発話発生時のゲーム内状況を、先ほど分類した発話音声データに対応付け仮想プレイヤーの分類済み発話音声セットとする。

発話音声データは、実際に格闘アクションゲームをプレイしている最中のプレイヤーの発話を録音し、

そこから感情的発話と判断できるものを抽出して作成する。音声は実際のプレイ中の一般人のものとした。同様にひとつのプレイ中は一人の発話音声のみを用い他の声にしない。これは音声データを単一の人物だけとした場合と複数人物の音声を切り替えて提示した印象評価実験から定めた。

収録は実際のゲームプレイ中の様子の撮影と、発話音声の録音とを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現した短文、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 前後の発話音声データが抽出できた。

この発話音声データ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。

4.2. システムの具体的構成

ゲームプレイ時点の動作は下記の流れとなる。このとき市販ゲーム機とゲームソフトが併用される。

【画面マーカ値抽出】

ゲームのキャプチャ映像から特定のゲーム内状況を表す特定の情報を検出しゲーム内状況を読み取る。

【プレイヤー心理パラメータ算出】

読み取ったゲーム内状況により、快-不快パラメータと発話頻度パラメータを求める。

【発話音声セット選択】

ゲーム内状況と、快-不快パラメータを参照して、対応する発話音声データを呼び出す。

【発話音声決定・発話】

音声として出力する。音声はゲーム内キャラクターの擬声語とは声色が異なるものとしたり、再生の場所を変えたりして、プレイしている者が容易に聞き分けられるようにする。ゲームプレイ時のシステムイメージを図 5 に示す。

4.3. ゲーム状況読み取り

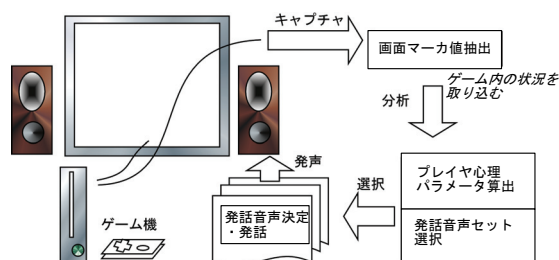


図 5 動作イメージ

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内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はゲームの内部プログラム情報を参照する必要が無いので数多くのゲームに応用できる点にある。また操作の結果のみを参照するのでコントローラの入力情報をも必要も無い。

具体的な方法としては図 6 のようにゲーム画面を PC で取り込み、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内状況を読み取る方法を用いた。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。この画素を閾値により 2 値化する、これの 20 個ずつの組み合わせによって 5% 分解能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。そして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。

大技はプレイヤーの感情に強く作用するので検出できることは重要である。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。これを利用しマーカの大部分が低い画素値を検出した場合を大技発生の兆候とみなす。その直後に被攻撃側がダメージを受けたなら大技の成功であると判定できる。或い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ることによっても大技成功の判断が可能である。

3.1.d)の「文章発話」が発生しやすい中休み状態の

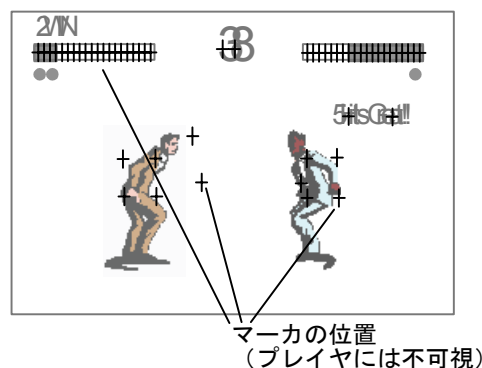


図 6 検出マーカの配置例

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、コマンド入力が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。

4. 4. ゲーム状況分析と発話の対応付け

読み取ったゲーム内状況と、3.1.で分類した発話音声データは次のようにして対応付ける。

a) 予期発話との対応付け

大技発生直前には画面の明度彩度が非常に低くなる。複数のマークよりその演出効果を読み取った場合に予期発話音声データを呼び出す。

b) 反応発話との対応付け

仮想対戦プレイヤー側の残りライフゲージ部分の画素値を読み取るマークを複数設置し、2フレーム前、1フレーム前両方との差分があればダメージを受けたと判断し、反応発話音声データを呼び出す。現在のマークの画素値と過去2フレーム分の画素値の差分を取ることで、キャラクター画像がライフゲージを覆う場合や、ノイズによって不正な画素値を検出した場合1フレームまでは対応する。

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、a)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。

c) 能動発話との対応付け

プレイヤー側の残りライフゲージ部分の画素値を読み取るマークを複数設置し、2フレーム前、1フレーム前両方との差分があればダメージを与えたと判断し、反応発話音声データを呼び出す。

人間は能動発話を厳密には攻撃を行う一瞬間に発声する場合が多いが、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、処理遅れについて問題はない。

d) 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両プレイヤー

のライフ差が大きいときに限定された発話内容などがある。これらはデータ1個単位でゲーム内状況との対応付けを行う。例えば試合終了の瞬間であれば、中央部分のマーク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、ある程度状況によらず発生する文章発話については、細かな対応付けはせず、上記a)b)c)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。

4. 5. 読み取り適正度の確認

以上の仕組みが正しく機能するかを確認する必要がある。実際に利用するゲーム等に依存する部分があるのでプレイ環境を構築してからこの確認を行う。

サンプルゲームとして、カプコン(株)の「STREET FIGHTER ZERO 2」を選定し、これを実際に数10分プレイさせ読み取りの正確さの検証を行った。その結果残りライフゲージなどの情報表示画面と検出された値とのタイムラグは0.06秒前後であり体感的遅延はなかった。

最も重要な残りライフポイントの値は、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。読み取れない場合は評価実験中に数回しか出現せず、無視できることがわかった。

大技の発生はほぼ全て検出できたが、その成否の検出は、技の発生からダメージ発生まで遅延が大きい場合や攻撃の一部のみのヒットの場合で検出もれが発生した。

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データが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得たものもあった。より細かなゲーム内状況の読み取りと、発話音声データの分類方法考案が今後の課題である。

4. 6. 心理パラメータのソフトウェア実装

3章で述べた感情モデルの概念式は多数の浮動小数点演算が含まれるため計算負荷が重いため以下の簡略化方式で実装した。

I. 残りライフポイント差に応じた発話頻度と心理パラメータの更新

単位時間のライフ値の変化に応じて表3のような変化を与える。

II. 残りライフに応じた発話頻度

表4のとおりとした。

III. 残り時間に応じた発話頻度の更新

表3 心理パラメータの更新ルール

ライフ差L	心理パラメータ更新	発話頻度更新
$75 \leq L \leq 100$	+0.05	-0.1
$50 \leq L < 75$	+0.033	-0.05
$25 \leq L < 50$	0	+0.05
$0 \leq L < 25$	0	+0.1
$-25 \leq L < 0$	0	+0.1
$-50 \leq L < -25$	-0.05	0
$-75 \leq L < -50$	-0.033	-0.05
$-100 \leq L \leq -50$	-0.05	-0.1

表4 残りライフ量と発話頻度更新量

ライフ差L	心理パラメータ更新
$75 \leq L \leq 100$	-0.1
$50 \leq L < 75$	-0.05
$25 \leq L < 50$	+0.05
$10 \leq L < 25$	+0.1
$0 \leq L < 10$	+0.2

時間単位 0.1 秒で 830 カウント (8.3 秒) に達した時点から+0.1 ずつ更新する。

5. 評価と考察

実装したシステムを実際にプレイしてもらい動作状況を確認し、その後意見を聴取した。

5.1. パラメータ推移の状況

本システムを適用して対戦アクションゲームをプレイした際の、パラメータの動きや発話音声データ再生の様子を例を図6に示す。この例では、快-不快パラメータの初期値は0、発話頻度パラメータの初期値は50とした。パラメータの各変動要因の重み付け量は、予備実験結果を見て経験的に設定した。

快-不快パラメータに関しては、大技発生時に大きく変動し、それ以外は残りライフポイントの変化に応じてゆるやかに推移しているのがわかる。発話頻度パラメータに関しては、時間経過と両者の残りライフポイントの平均の現象によって、後半に向かって緩やかに上昇しつつも、後半で人間のプレイヤーと仮想対戦プレイヤーの残りライフポイントの差が大きく開き、勝負がほとんど決してしまうと急激に値を減少させている。またパラメータは両方とも、後半に残りライフポイントが少なくなるにつれて、変動の様子が激しくなっている。これはウェーバー・フェヒナーの法則が正常に働いていることを示している。このグラフから4章で設定したパラメータ変動の式は、意図どおり働いていることが確認できた。

5.2. 実プレイの様子とアンケート

本システムを適用した格闘ゲームを大学生男女15名の被験者に体験させ、表5のアンケートをとり図

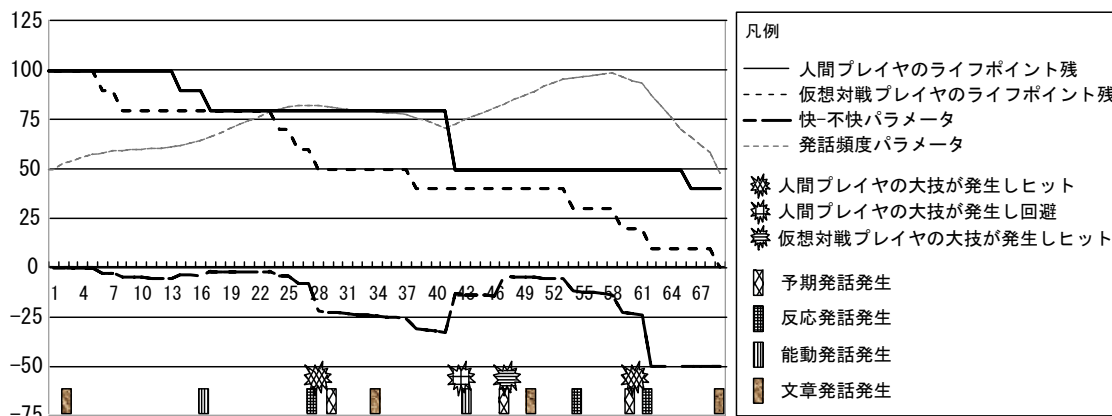


図6 パラメータ値の推移

表5 アンケート項目

- 1)COM(仮想対戦相手)が発話するのは楽しい
- 2)発話なしに比較してゲームを面白く感じる
- 3)発話なし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

回答選択肢

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

7の集計と表6の自由記入形式の感想を得た。

アンケート結果から、仮想対戦プレイヤはプレイヤ1人で対戦型ゲームを行う際の面白さの改善に寄与していることがわかる。

特に1)から5)の質問に対しては、音声データ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であったことが挙げられる。特に質問7)に対しては、面識のない被験者音声のほうが高評価である。この結果から、発話音声データ抽出元のプレイヤと知人友人関係であるプレイヤのほうがおもしろく感じていることが分かる。

6. まとめと今後の予定

本稿では、ゲーム映像からゲーム内状況を判断し快-不快と没入度の内部パラメータを変化させゲーム

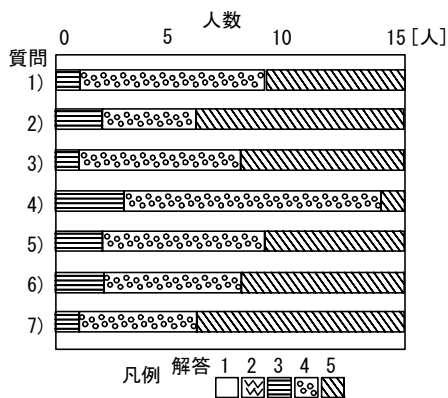


図7 アンケート結果

表6 自由回答結果

- ・状況判断ができてよい。
- ・タイミングよく反応が返りストレスを感じない
- ・声でもとの人物と対戦しているような錯覚を得た
- ・COM が喜ぶ声に発奮し次回勝利への意欲が出た
- ・感情設定の存在が判り変化に応じた発声が面白い
- ・COM が感情的発話をするついで笑ってしまう
- ・状況にそぐわない発話があり、その場合は不快
- ・知人の声の実装されたものでもやってみたい

内状況とパラメータに対応付けした発話音声データ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を生成する手法を提案した。さらに市販の格闘アクションゲームプレイ環境に適用し当手法は対戦型ゲームをプレイヤ1人でプレイする場合の面白さの向上に寄与することがわかった。

今後の課題として人対人との発話内容の客観的な評価方法の検討と実施がある。また多人数の発話状態をサンプリングして統計分析しより完成度の高いモデルの構築を行う。また対象ゲームを本実施例の様な幾分古典的内容だけでなくより高度な空間的動きを伴うもの、グループ戦を行うもの、オンラインゲームなどへの適用可否などを含み拡大する方法についての検討を進める。現在の対象ゲームについても発話音声データの作成および分類の手作業の一部自動化を目指す必要がある。特に身近な知人の発声を手軽にシステムに取り込めるような手法が確立されれば応用性も高まる。また仮想対戦プレイヤ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データの内容やタイミングならプレイヤが面白いと感じるか等の心理的検証を深める。

参考文献

- [1] C.Breazeal and B. Scassellati: "A context-dependent attention system for a social robot." Proceedings of the International Joint Conference on Artificial Intelligence, Morgan Kaufmann Publishers, pp. 1146-1151, 1999
- [2] Takanishi, A. Sat, K. Segawa, K. Takanobu, and H. Miwa, H.: "An anthropomorphic head-eye robot expressing emotins based on equations of emotion," IEEE International Conference on Robotics and Automation(ICRA2000), 2000
- [3] Hiroki Ogawa , Tomio Watanabe.: "InterRobot: speech-driven embodied interaction robot", Advanced Robotics, Vol.15, No.3, pp.371-377, 2001.
- [4] 中島宏, 森島泰則, 山田亮太, Scott Brave, Heidy Maldonado, Clifford Nass, 川路茂保.: "人間-機械協調システムにおける社会的知性 ー心のモデルとパーソナリティーによるエージェントの社会的応答についてー", 人工知能学会論文誌, 19 卷 3 号 C, 2004
- [5] 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.
- [6] 深山篤, Vincent Bao 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] 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.
- [8] 千葉浩彦, 佐伯胖, 佐々木正人編.: "アクティブ・マインド", 東京大学出版会, 1990.
- [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.