

■S3 群 (脳・知能・人間) - 感性・マルチモーダル情報処理

4 章 感情を伴い気遣う仮想インタフェースエージェント

(執筆著者：ヘルムト・ブレンディンガー，石塚 満) [2009年4月 受領]

■概要■

仮想エージェントあるいは生命的エージェントは、合成音声、顔表情、身振り、会話における視線の動きなどを用いて対面コミュニケーションの幾つかの側面を模擬する、一定レベルの自律性を有する擬人化キャラクタである^{1), 2)}。マルチモーダルな身体的行動を通して、感情や関心事を表出するのも、人間的親しみやすさを与える点で重要となる。また、最近のエージェントは、ユーザの感情や関心事を検知する機能ももつようになってきている。このように、ユーザの状態の検知や感情表出などにより、エージェントはユーザと信頼感や感情を伴うコミュニケーションを達成する。

ここでは、感情を伴うインタラクティブゲーム、関心に基づくプレゼンテーションを例にして、そのようなインタフェースエージェントについて紹介する。ユーザの感情状態は、皮膚導電率や筋電位 (EMG) 信号のような生体情報から、継続的に取得することができる。これにより、例えばユーザのフラストレーション状態を検出し共感の情を示したり、ユーザの感情に反応して気配りするなどのエージェントを構築できることになる。また、目の動きや視線方向からユーザの関心事を推測することができる。

これらの機能は、ユーザとのインタラクションの質の向上に寄与すると考えられ、多くの研究が³⁾⁻⁸⁾ が成されてきているものの、どの程度効果的であるのかは十分に知られていない。なかには効果に関して否定的な面を示している研究例もある⁹⁾⁻¹¹⁾。また、アンケートによる古典的評価法が人間-コンピュータ間の感性的インタラクションの評価に適切かどうかを問う研究もある^{12), 13)}。ここでの主要な論点は、実験後の自己報告が瞬時瞬時のインタラクション体験の評価として十分でなく、必ずしも信頼できないという点である¹⁴⁾。

生命的エージェントとの感性的インタラクション評価の別のアプローチは、精神生理に基づく方法の使用であり、精神生理現象からは人間の感情状態に関する良い情報が得られることが知られている¹⁵⁾。このアプローチの利点は以下になる。

- ・ユーザのインタラクション体験の動的性質を推測できる。
- ・精神生理は意図的には制御できないので、虚偽の報告を防止できる。
- ・ユーザの認識の判断や記憶能力に依存しない。
- ・インタラクションの主なタスクに干渉されない。

一方、注視は人間の意図的制御によるものであるが、眼や瞳の動きは意図的制御を超えて、環境への人間の解釈、反応を表す良い情報を提供する¹⁶⁾。本章では、生命的エージェントインタフェースにおける精神生理信号と視線情報の利用の可能性について記す。

■感性・マルチモーダル情報処理 - 4章

4-1 感情を伴うインタフェース

(執筆者：ヘルムト・ブレンディンガー，石塚 満) [2009年4月 受領]

感情をもつエージェントの効果についての研究は多くはないが、Braveら¹⁷⁾はブラックジャック・ゲームにおける検討を行っている。このカジノ形態のインタラクションにおいては、人間の顔写真をもつプレーヤがおり、次の4種の感情的反応を表出する。

- 1) 自己中心的（勝てば喜び，負ければ悲しむ）
- 2) 共感的（勝者に対して喜び，敗者に対して気の毒に思う）
- 3) 自己中心的と共感的の両者
- 4) 両者とも無し

反応に付随してエージェントに関連するテキストボックスに、例えば共感的の場合“*That's great! I'm really happy that you won.*”のように表示される。アンケート法の評価では、被験者は共感を示すエージェントをより好ましく信頼できると評価し、思いやりがあると感じている。

このBraveらの研究は感情を伴うエージェントの効用を支持するものではあるが、以下のような限界もある。ユーザがエージェントとインタラクションする状況はディーラと対峙するカジノの状況よりも日常的なものであること、生命的エージェントはここでの写真顔のエージェントよりも豊かな表現モダリティをもつこと、そしてアンケート方式の評価ではユーザの瞬時瞬時の体験を十分に評価できないことである。

Prendingerら¹³⁾は、クイズゲームにおいて意図的に誘発されたフラストレーションが共感的（謝罪する）エージェントにより軽減されることを、皮膚導電率の計測によって示している。

4-1-1 感情モデル

ここでは、感情モデリングと認識についての二つの影響力をもつ理論について記す。第一は、OCCモデル¹⁸⁾として知られるOrtony, Clore, Collins (OCCはこの3者の頭文字)による感情モデルである。第二は、興奮度と正負度 (Valence) の2軸による座標空間で感情を表すLangのモデル¹⁹⁾である。

OCCモデルでは感情は、他の動作主体の行動や事物の感じ方を含むイベントへの、自分の目標やスタンダード、嗜好を反映した正/負の反応として捉えられる。このモデルにおける事象の重要度は、以下の4種の抽象化された心的概念への関係を含む感情誘起条件により決定される。

- ・信念 (Belief)：主体が成立しているとの証拠を有する事象の状態
- ・目標 (Goal)：主体にとって好ましい（あるいは好ましくない）事象の状態，即ち得たいと欲する（欲しない）こと。
- ・スタンダード：あるべきこと（あるべきでないこと）に関する主体の信念，即ち称賛すべき（非難すべき）と考えられるイベント
- ・態度 (Attitude)：他の主体やオブジェクトに対する好き・嫌いに関する主体の性向

OCC感情モデルによれば、感情の型は言葉でラベル付けされた誘起条件のクラスとなる。

計 22 種の感情が特定されており，“喜び (Joy)”、“苦痛 (Distress)”、“他者への喜び (Happy for)”、“満足 (Gloat)”、“不快 (Resent)”、“遺憾 (Sorry for)”、“非難 (Reproach)”、“怒り (Angry at)” などである。単純な感情の例は“喜び”であり、主体がある目標をもち、それが達成された際に喜びを体験することになる。目標がどの程度好ましいかに依存して、感情強度は変化する。より複雑な感情は“怒り”であり、1)主体 B が動作 a を行い、それが主体 A の目標に相反する状況を生じさせ、2)A が動作 a を非難すべきこととみなす場合に、A は B に怒りを感じるようになる。

思いやりのあるユーザインタフェースの設計にはユーザの感情モデリングを必要とする。他者の幸福である“他者への喜び (Happy for)”感情について考えると、1)主体 A が主体 B を好ましく思い、2)B が嬉しいときに、A は B に対して喜びの感情をもつ。このタイプの感情である“他者への喜び (Happy for)”、“満足 (Gloat)”、“不快 (Resent)”、“遺憾 (Sorry for)”は他者の感情状態の検知という難しい課題をもたらす。この課題に対処するには少なくとも二つのアプローチが存在する。第一は典型の利用であり、例えばコンピュータゲームのプレイヤーは勝つ目標をもち、勝つことにより嬉しくなると仮定するなど、ユーザは典型に分類されると仮定する²⁰⁾。第二は、顔表情、音声の韻律、言語的スタイル、動作等などのコミュニケーションモダリティ²¹⁾から推測される、ユーザの感情表出に基づくアプローチである。

新しいアプローチとしては、ユーザ感情を認識するのに精神生理信号を用いることがある。Lang¹⁹⁾や Feldman-Barnett, Russel²²⁾により提唱された 2-D 感情モデルでは、すべての感情は次の 2 軸によって特徴付けられるとされている。

- ・ 正負度 (Valence) : 快適か不快か
- ・ 興奮度 (Arousal) : 冷静か興奮か

ここで、名前付けられた感情名はこの 2 軸による空間の点として表される。例えば、“悲しみ (Sadness)”は低興奮度で負の感情、“怒り (Anger)”は高興奮度で負の感情、“幸福感 (Happiness)”は低中の興奮度で正の感情である。

自律神経システムと上記 2 軸の関係は精神生理学の研究¹⁵⁾に基づいている。以下には例として、三つの重要な精神生理信号である皮膚電気反応 (GSR)、筋電図 (EMG)、容積脈波 (BVP) についての機能について記す。

- ・ 皮膚電気反応 (GSR) : GSR 信号は皮膚導電率の指標であり、皮膚腺からのイオンを含む汗が導電率を変化させ、これを指につけた電極を通して測定する。皮膚導電率は個人の全体的興奮度レベルに比例して増大する。
- ・ 筋電図 (EMG) : EMG 信号では、微小筋繊維が収縮する際に発生する表面電圧を (通常、下肢または咀嚼筋部で) 検出することにより、筋活性度を測定する。平均筋活性度は、感情の正負度と負の相関をもつ。
- ・ 容積脈波 (BVP) : BVP 信号では、皮膚に赤外線光を当て、どの程度反射されるかを測定することにより、血流量を測定する。これにより心拍数と心拍間隔も測定できる。心拍数は不安や恐怖のような負の感情に伴って増加する。

4-1-2 インタラクティブゲーム

感情を伴う対面型インタラクティブゲームの例として、カードゲーム“Skip-Bo”の例について紹介する²³⁾。このゲームでは、プレイヤーはプレイを通じてテーブル右側の個人の山の 8

枚のカードを白い中央共有スタックへ移すという目標をもつ(図 4・1 参照)。中央スタックではカードに順番(1 から 13 まであり)があるので、目標を達成し勝利するために、手持ちと個人の山のカードを戦略的に使用しなければならない。MAX と呼ばれるのがここでの対戦相手の仮想エージェントであり、音声出力、顔表情と身体ジェスチャなどのマルチモーダルなインタラクション機能を有する。例えば、感情状態に応じた幸福度-興奮度-支配度(PAD: Pleasure-Arousal-Dominance)空間内での、幾つかの顔表情を表出することができる²⁴⁾。また、低い呻き声のような感情的音声を出したり、継続的に呼吸や瞬きも行う。また、自分の手が終了すると、頷く合図により相手に手番を渡す。

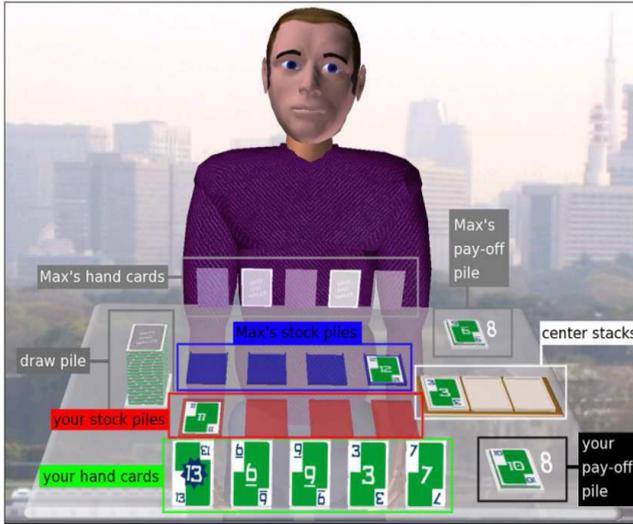


図 4・1 仮想カードゲーム“Skip-Bo”の環境 - 正面は対戦相手の MAX エージェント²³⁾

共感人間-コンピュータ間のインタラクションで重要な側面となることが知られている^{17), 25), 26)}。ゲームにおいては、共感ユーザ感情へのエージェントの反応を指し、正の反応(ユーザの苦悩を気の毒に思う)と負の反応(ユーザの苦悩を嬉しく思う)の両者を含む。Skip-Bo は競合ゲームであり、ユーザはエージェントを敵とみなしている。ゆえに、敵である仮想エージェントがその目標に従って正と負の感情を伴う動作を表出するように自然に振る舞う場合に対し、もしエージェントがそうしない場合にはユーザの興奮度とストレスは低くなると想定できる。この効果を検証するために、以下のシナリオに従った実験が行われた²³⁾。

- ・無感情条件：エージェントは何の感情的振る舞いをしない。
- ・自己中心的感情条件：エージェントは、例えばカードを移せるときに嬉しい顔をするなど、自己プレイのみ賞賛する。
- ・負の共感条件：エージェントは自己中心的感情動作をし、そしてユーザに負の反応を示す。またユーザが良いカード移動を行うときに苦悩を表出し、ユーザが悪いカード移

動で苦悩が認められるときに喜びを表出する。

- ・ **正の共感条件**：エージェントは自己中心的感情動作をし、かつユーザの動作を正に評価し、その結果、ユーザはゲーム進行が嬉しくなる。即ち、もしユーザが苦悩していると感じられたらエージェントは気の毒に思い、悲しみを表出する。

実験では 32 名の被験者に対して皮膚電導率と筋電図 (EMG) が計測された。無感情と正の共感条件の場合、エージェントが勝利のカード移動を行うときの皮膚電導率から、人間プレーヤーは最も興奮度 (またはストレス) が高くなることが示された。この結果は正の共感反応に対して直観に反するのように感じられるが、競合カードゲームで正の共感 (または無感情) を見せることは大変不自然に映るためである。同様な結果はユーザが勝利のカード移動する場合の皮膚電導率にも観測され、負の共感条件時よりも正の共感条件時により大きな興奮度 (またはストレス) が体験される。

エージェント及びユーザの勝利のカード移動時の筋電図 (EMG) の結果は、正/負の共感条件時に大きな差異が現れている。例えば、エージェント側の負の感情はユーザ側の負の感情を誘発するといったように、筋電図はエージェントの感情の正負度がユーザの精神生理レベルで反映し、相互依存することを示す。このように競合ゲームでの敵の役割のエージェントは敵として期待されるように振る舞うべきであることが言え、もしエージェントが教師の役割を果たす教育インタフェースとしてゲームが設定される場合には、実際の教師のようにエージェントが正の共感のフィードバックを示すときに、ユーザはストレスを受けることが少ないと推測されることになる。

■感性・マルチモーダル情報処理 - 4章

4-2 気遣うインタフェース

(執筆者：ヘルムト・ブレンディンガー，石塚 満) [2009年4月 受領]

気遣うユーザインタフェース (AUI : Attentive User Interface) ²⁷⁾，または気遣う視覚的インタフェース ²⁸⁾ では，ユーザの視線がユーザ動作の状況情報を提供するインタラクション・スタイルが構想されている。Vertegaal ²⁹⁾ では，「気遣うインタフェースでは，ユーザとシステム両者の情報処理能力が複数タスクに最適に分配されるように，ユーザに提示する情報を動的に優先付ける。このインタフェースはこの優先付けを観測とユーザ関心事の過去・現在・未来のモデルとの組合せから成る知識に基づいて行う」としている。Starke と Bolt ³⁰⁾ により記された視線反応型自動開示ディスプレイは，ユーザの視線を監視し，視線動作パターンに実時間で反応する初期のシステム例である。視線はまたユーザの関心事にマルチメディアシステムを自動的に適合させる上で中心的な役割を果たし，いわゆる“関心と感情に反応するメディア (Interest And Emotion Sensitive Media)” ³¹⁾ の実現に役立つ。

4-2-1 視線からの関心事の推定

ここでは視線動作からユーザの関心事を定めることを可能にする，Qvafort と Zhai ³²⁾ によるアルゴリズムを簡単に記す。

- ・関心度スコア (IScore) は事物への関心のレベルを把握する。もし IScore がある閾値を超えると，その事物は活性化したと言われる。
- ・関心度スコアの焦点 (FIScore) は活性化事物への関心度を時間軸に沿って計算する。もし FIScore がある閾値以下になると，その事物は非活性となり，ユーザは興味を失ったとみなされる。

IScore の基本要素は $p = \text{TISon} / \text{TIS}$ であり，ここで TISon は TIS (1000 ミリ秒に設定される) の時間窓内での視線が向けられた留積時間である。興味を増大，あるいは減少させる要因を説明するために，IScore は以下のように定められている。

$$\text{Pis} = p (1 + \alpha (1 - p))$$

ここで， α は関心度推定精度を向上させるパラメータ集合を記号化したものである。例えば，以下 4-2-2 項に記す気遣うプレゼンテーションエージェントの例では，簡単化したアルゴリズムが適用されており，次の二つのパラメータ α_f と α_s が用いられている。この α_f は事物への興味を表すことになる事物への視線の入出の回数であり， α_s は対象事物のサイズに比較した際の全関心対象事物の平均サイズを表す。 α_s は関心対象物のサイズの違いを補正するものである。FIScore は活性事物へのユーザの連続的関心度を計測するものであり，IScore と同様にその基本要素は事物へのユーザ視線の強度である。FIScore は時間窓 (2000 ミリ秒に設定される) での他の事物への視線強度も考慮した補正が行われる。

4-2-2 気遣いプレゼンテーションエージェント

図 4・2 に示すのは，視線情報を利用する気遣いプレゼンテーションエージェントを例示している ³³⁾。視線情報を利用してユーザの関心事を把握することにより，ユーザ個人に適合するプレゼンテーションを生成している。ここではユーザの視線情報の次の 2 種のタイプに応

じた反応が実装されている。

- ・スクリーン上のオブジェクトに対する関心，非関心に対応する文脈非依存のエージェントの反応（この割り込みはプレゼンテーション中の任意の時点で発生するので，文脈非依存である）。
- ・失敗したグラウンディング状況に対する文脈依存のエージェントの反応（人間の対面コミュニケーションにおいて，グラウンディングは話したことが対話相手に理解されたとの確証を得るプロセスに関係しており，理解されれば共通グラウンド³⁴⁾が存在することになる）。



図 4・2 赤外線による視線追跡のステレオ装置をもつ気遣いプレゼンテーションエージェントのセットアップ³³⁾

後者の場合については，例えば話題の事物にユーザの関心が向けられていない場合に共通グラウンドの確立に向け，プレゼンテーションエージェントは話す内容とスクリーン上の物との間に参照関係を確立させようと，指示動作を繰り返したりすることになる。

4-2-3 むすび

本章では，精神生理信号を用いる感情を伴うインタフェースと，視線情報を用いる気遣うインタフェースを中心に紹介した。それらの効用の評価，検証も進みつつある。今後，計測デバイスの小型化などの進歩につれて，実用性は高まるであろう。

■参考文献

- 1) J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, “Embodied Conversational Agents,” Cambridge, MA: MIT Press, 2000.

- 2) H. Prendinger and M. Ishizuka, "Life-Life Characters. Tools, Affective Functions, and Applications," Berlin, Heidelberg: Springer, 2004.
- 3) J. Bates, "The role of emotion in believable agents," *Communications of the ACM*, **37**(7), 122-125, 1994.
- 4) E. André, M. Klesen, P. Gebhard, S. Allen, and T. Rist, "Integrating models of personality and emotions into lifelike characters," *Proceedings International Workshop on Affect in Interactions—Towards a New Generation of Interfaces*, pp.136-149, 1999.
- 5) A. Paiva (Ed.), "Affective Interactions: Towards a New Generation of Computer Interfaces," LNAI 1814, Berlin: Springer, 2000.
- 6) H. Prendinger, S. Descamps, and M. Ishizuka, "Scripting affective communication with life-like characters in web-based interaction systems," *International Journal of Applied Artificial Intelligence*, **16**(7-8), 519-553, 2002.
- 7) A. Ortony, "On making believable emotional agents believable," In P. P. R. Trappl, "Emotions in Humans and Artifacts," pp.189-211, Cambridge, MA: MIT Press, 2003.
- 8) C. Pelachaud and M. Bilvi, "Computational model of believable conversational agents," *Communication in Multiagent Systems: Background, Current Trends and Future*, pp.300-317, Berlin Heidelberg: Springer, 2003.
- 9) J. C. Lester, S. A. Converse, S. E. Kahler, S. T. Barlow, B. A. Stone, and R. S. Bhogal, "The persona effect: Affective impact of animated pedagogical agents," *CHI'97*, pp.359-366, ACM Press, 1997.
- 10) D. M. Dehn and S. van. Mulken, "The impact of animated interface agents: A review of empirical research," *International Journal of Human-Computer Studies*, **52**, 1-22, 2000.
- 11) D. C. Berry, L. T. Butler, and F. de. Rosis, "Evaluating a realistic agent in an advice-giving task," *International Journal of Human-Computer Studies*, **63**(3), 304-324, 2005.
- 12) R. Picard, S. B. Daily, "Evaluating affective interactions: Alternatives to asking what users feel," *CHI'05 Workshop on Evaluating Affective Interfaces: Innovative Approaches*, 2005.
- 13) H. Prendinger, J. Mori, and M. Ishizuka, "Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game," *International Journal of Human-Computer Studies*, **62**(2), 231-245, 2005.
- 14) R. E. Nisbett and T. D. Wilson, "Telling more than we know: Verbal reports on mental processes," *Psychological Review*, **84**, 231-259, 1977.
- 15) J. L. Andreassi, "Psychophysiology: Human Behavior & Physiological Response," Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- 16) A. T. Duchowski, "Eye Tracking Methodology: Theory and Practice," London: Springer, 2003.
- 17) S. Brave, C. Nass, and K. Hutchinson, "Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agents," *International Journal of Human-Computer Studies*, **62**(2), 161-178, 2005.
- 18) A. Ortony, G. L. Clore, and A. Collins, "The Cognitive Structure of Emotions," Cambridge, MA: Cambridge University Press, 1988.
- 19) P. J. Lang, "The emotion probe: Studies of motivation and attention," *American Psychologist*, **50**(5), 372-385, 1995.
- 20) E. Rich, "User modeling via stereotypes," *Cognitive Science*, **3**, 329-354, 1979.
- 21) R. Picard, "Affective Computing," Cambridge, MA: MIT Press, 1997.
- 22) L. Feldman-Barrett and J. A. Russell, "The structure of current affect: Controversies and emerging consensus," *Current Directions in Psychological Science*, **8**(1), 10-14, 1999.
- 23) H. Prendinger, C. Becker, and M. Ishizuka, "A study in users' physiological response to an emphatic interfaces agent," *International Journal of Humanoid Robotics*, **3**(3), 371-391, 2006.
- 24) C. Becker, S. Kopp, and I. Wachsmuth, "Simulating the emotion dynamics of a multimodal conversational agent," *Proceedings Tutorial and Research Workshop on Affective Dialogue Systems (ADS'04)*, pp.154-165, Berlin: Springer, 2004.
- 25) J. Klein, Y. Moon, and R. Picard, "This computer responds to user frustration: Theory, design, and results," *Interacting with Computers*, **14**, 119-140, 2002.
- 26) A. Paiva, J. Dias, D. Sobral, R. Aylett, P. Sobreperez, S. Woods, C. Zoll, and L. Hall, "Caring for agents and agents that care: Building emphatic relations with synthetic agents," *Proceedings Third International*

- Conference on Autonomous Agents and Multi-Agent Systems, pp.194-201, New York: ACM Press, 2004.
- 27) S. Zhai, "What's in the eyes for attentive input," *Communications of the ACM*, **46**(3), 34-39, 2003.
 - 28) T. Selker, "Visual attentive interfaces," *BT Technology*, **22**(4), 146-150, 2004.
 - 29) R. Vertegaal, "Designing attentive interfaces," *Proceedings of the Symposium on Eye Tracking Research & Applications (ETRA'02)*, pp.22-30, New York: ACM Press, 2002.
 - 30) I. Starker and R. A. Bolt, "A gaze-responsive self-disclosing display," *Proceeding of CHI'90*, pp.3-9, New York: ACM Press, 1990.
 - 31) J. P. Hansen, A. W. Andersen, and P. Roed, "Eye-gaze control of multimedia systems," *Proceedings 6th International Conference on Human Computer Interaction (HCI'05)*, pp.151-190, Elsevier, 1995.
 - 32) P. Qvarford and S. Zhai, "Conversing with the user based on eye-movement patterns," *Proceedings of CHI'05*, pp.221-230, New York: ACM Press, 2005.
 - 33) T. Eichner, H. Prendinger, E. André, and M. Ishizuka, "Attentive presentation agents," *Proceedings of 7th International Conference on Intelligent Virtual Agents (IVA'07)*, pp.283-295, Berlin, Heidelberg: Springer, 2007.
 - 34) H. H. Clark and S. E. Brennan, "Grounding in communication," In J. L. L.B. Resnick, "Perspectives on Socially Shared Cognition," pp.127-149, Washington: APA Books, 1991.