

人間と擬人化エージェントによる感情マッピングの相互学習

Acquiring emotion mappings through the interaction between a user and a life-like agent

山田 誠二

Seiji Yamada

東京工業大学

大学院総合理工学研究科知能システム科学専攻

CISS, IGSSE, Tokyo Institute of Technology

山口智浩

Tomohiro Yamaguchi

奈良工業高等専門学校 情報工学科

Department of Information Science

Nara National College of Technology

Abstract: This paper describes a human-agent interaction framework in which a user and a life-like agent mutually acquire their emotion mappings through a mutual mind reading game. In these several years, a lot of studies have been done on a life-like agent like a Micro Soft agent, an interface agent, and so on. Through the development of various life-like agents, emotion has been recognized to play an important role in making them believable to a user. For making effective and natural communication between a life-like agent and a human user, they need to identify the other's emotion state from expressions and we call mappings from expressions to emotions *emotion mappings*. If an agent and a user don't obtain these emotion mappings, they can not utilize behaviors which significantly depend on the other's emotion states. We try to formalize the emotion mapping and a human-agent interaction framework in which a user and a life-like agent mutually acquire emotion mappings each other. In our framework, a user plays a mutual mind reading game with a life-like agent and they gradually learn emotion mappings each other through the game.

1 はじめに

近年、人や動物と同じような外見をもち、ユーザに対する支援や情報提示などのインタラクションをもつソフトウェアエージェントである擬人化エージェント (like-like agent) の研究が活発にされている [5]。その中でも、感情モデルの研究を基に擬人化エージェントに感情を持たせて感情表現を生成することにより、ユーザとエージェント間のコミュニケーションにおける信頼性や自然さを高めようとする研究が始まっている [1]。

このような擬人化エージェントとユーザがインタラクションをもつとき、ユーザはエージェントの表情などを基にエージェントの感情を初めとする内部状態を推定するようになる。反対に、エージェントがユーザの表情などからユーザの感情などの心理的状态を推定しようとする。なぜなら、お互いが相手の感情を把握することにより、相手と適切なインタラクションを構成できるからである。このようなことは、相手の感情を考慮しながらコミュニケーションしている人間にとっては自然なことであり、擬人化エージェントとユーザが自然にコミュニケーションする意味で重要である。

エージェントに感情を持たせる研究では、エージェントは人間のユーザの感情を観測により知覚でき、またユーザも擬人化エージェントが生成した表情などを元にエージェントの感情を同定できるという仮定のもとにシステムが設計されている。しかし、このような仮

定は、実際の人間の表情をできる限り再現する擬人化エージェント [3] であれば成り立つ可能性もあるが、人間に似た顔の表情をもたないエージェントなどを含む一般の擬人化エージェントでは成り立ち難い。つまり、領域依存の知識をうまく入れることなくしては、エージェントはユーザの表情からユーザの感情を同定できないし、逆にユーザもエージェントの表情からその感情を同定できるものではない。また、感情-表情間の対応関係には、個人差、民族差、時代や歴史性などによる差異やバラツキが考えられ、不特定のユーザに対して唯一の対応関係を仮定することは無理がある。

そこで本研究では、ユーザと擬人化エージェント間が相互読心ゲームという表情に基づくコミュニケーションを通じて、互いが相手の表情と感情の対応である感情マッピングを学習していく枠組みを提案する。

本研究は、ユーザとエージェントがインタラクションをもつシステムでは、両者の間のコミュニケーションを円滑にするために、必ずユーザ側でエージェントに対する適応が起こり、逆にエージェント側でユーザに対する適応も必要になって、ユーザとエージェントの相互適応が起るという考え方である人間とエージェントの相互適応 HAMA (Human-Agent Mutual Adaptation)¹を基盤にしている。ここで、我々は、これまでのように機械学習や適応アルゴリズムを単体で研究するのではなく、HAMA という枠組み全体が上手く機能するよ

¹<http://www.ymd.dis.titech.ac.jp/hama/>

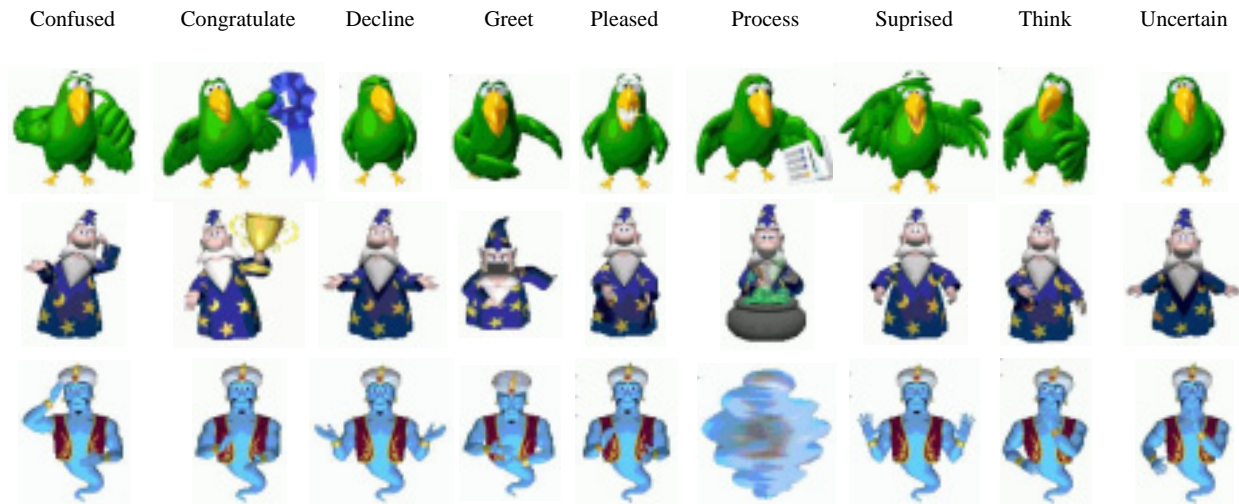


図1 MS エージェントの感情と表情

うに、特にユーザとエージェント間のインタラクションの設計に重点をおいて研究を行なう。そして、相互適応が円滑に進むための、ユーザとエージェント間のインタラクションの具体的方法として、お互いが相手の感情マッピングを学習することが目的である協動的相互読心ゲームを提案する。

Velásquez は、心の社会に基づいた感情生成モデル Cathexis を提案している [15][12][13][14]。Cathexis は心の社会を基本的枠組みとし、基本的な感情を司るプロトスペシャリストエージェントの相互作用により、感情の顕在化のみならず、ムードや気質の生成も可能なモデルとなっている。また、ロボットやソフトウェアエージェントを用いた実験により、自然な感情生成が実現できることを検証している。しかし、Cathexis は感情生成のモデルであり、表現と感情との対応を扱っているわけではない。

飯田と原は、人間の教示特性が、それまで継続的に間違っていた行為選択が正しい選択を行った場合に特に多くの報酬を与える傾向があることを考慮した強化学習を用いて顔ロボットの表現を学習させている [4]。しかし、感情状態と表情の対応を学習することは行っていない。

角所ら [9] は、ユーザ同士が相手の顔の CG 画像を見ながらコミュニケーションするエージェントメディアエテッドコミュニケーションにおいて、ユーザが本当に表出したい表情を理想表情とし、その理想表情と現実生成された表情との間のギャップを事例に基づく学習により補正する研究を行なっている。この研究の目的は、本研究の目的の表情-感情の相互学習とは異なるが、ユーザ同士の相互学習が可能な枠組みであり、本研究と関連が深い。

また、人間の顔の表情の認識は、コンピュータビジョンを用いたさまざまなアプローチの研究 [7][6] が行な

われている。これらの技術は、本研究において、エージェントがユーザの表情を認識するための要素技術として利用できる。

人間とエージェントではなく、エージェント同士が相互に語彙を獲得する研究がある [10][11][16]。語彙の獲得とは、ある物理的特徴をもった指示対象に名前(文字列)を割り当てていくことを意味する。これらの研究は、必ずしもエージェント間で同一の特徴に同一の名前が割り当てられる必要はないこと、また名前は指示対象とそれ以外の対象が区別できる特徴に割り当てられるという区別ゲームの考えに基づいている。本研究との共通点もあるが、エージェントと人間のインタラクションであること、同一指示対象ではなく、それぞれが相手から表出された表情を指示することなどが異なる。

2 表情からの感情の同定

現在さまざまな擬人化エージェントが開発されているが、それらは、さまざまな表情をもっている。テキストや音声によるコミュニケーションに加えて、表情によって感情などの内部状態を相手に伝達することは、人間どおしのコミュニケーションにおいては自然に行なわれていることであり、人間との自然なコミュニケーションを目指す擬人化エージェントも表情によるコミュニケーションをユーザと行なうことが重要である。また、エージェントは、ユーザの感情の状態を知ることにより適切な行動選択も可能になる。

しかし、前述のように、感情-表情間の対応関係には、個人差、民族差、時代や歴史性などによる差異やバラツキが考えられ、不特定のユーザに対して唯一の対応関係を仮定することはできない。このため、どのようなユーザでも唯一に解釈できるような表情の生成は容易

ではない。例えば、現在もっとも普及しているであろう擬人化エージェントの一つに、MicroSoft Agent²がある。それを基に開発された擬人化エージェントのアプリケーションにおいて、感情などの内部状態が実際にどのような表情により表現されているかの例を図1に示す。この図は、WebSpeak³というMicroSoft Agentが動くWebページの作成支援アプリケーションで用いられている3種類のエージェントPeedy, Merlin, Genieのそれぞれの感情などの内部状態(confused, congratulateなど)とその時の表情を示している。この図から、その表情から感情などの内部状態を一意に決定することは容易ではないことがわかるだろう。また、最も開発に労力が投入されているであろうMicroSoft Agentでさえも感情の同定は容易ではないことから、他の種々の擬人化エージェントではさらにエージェントの感情の同定は困難であると考えられる。

また、逆にエージェントが一般のユーザの表情からその感情を同定することも、個人差や不安定性のため容易ではない。よって、ユーザと擬人化エージェントが必要に応じて、お互いの表情からの感情同定を学習する必要がある。

3 感情マッピング

表情-感情の対応をユーザと擬人化エージェントが相互学習するための枠組みを与えるために、まず以下の概念を定義する。

- 感情状態 (emotion state) S_i : エージェント i の感情の状態を表す変数が S_i であり、これに次の基本感情が代入される。
- 基本感情 (primitive emotion) $E_i = \{e_1^i, \dots, e_m^i\}$: 感情のプリミティブのこと。 E_i は、エージェント i のもつ基本感情の集合である。心理学で提案されているものを使う。また、感情の生成モデル [12] で扱っている、ムードとか複雑なものは、とりあえず扱わない。
- 基本表情 (primitive expression) $X_i = \{x_1^i, \dots, x_n^i\}$: エージェント i のもつ基本的な表情。現状では、複数の基本表情の組み合わせから表出される表情は考えていない。
- 感情マッピング (emotion mapping) $M_{i:x \rightarrow e} = \{x_1 \Rightarrow e_1, \dots, x_m \Rightarrow e_m\}$: エージェント i の表情から感情への1対1写像。つまり、ひとつの表情は唯一の基本感情に写像される。また、エージェント j が学習したエージェント i の感情マッピングを上付きのサフィックス j を付けた $M_{i:x \rightarrow e}^j$ で表す。

- 表情マッピング (expression mapping) $M_{i:e \rightarrow x} = \{e_1 \Rightarrow x_1, \dots, e_m \Rightarrow x_m\}$: エージェント i の感情から表情への1対1写像。
- 感情遷移規則 (emotion transition rule) $T^a = \{e_i^a \times e_b \Rightarrow e_j\}$: 基本感情 e_i の状態にあるエージェント a が、相手エージェント b の感情を e_b と同定した場合に、エージェント a が次に遷移する基本感情 e_j を決める規則しかし、(決定論的な)感情遷移規則があるという仮定は難しい。時間やコンテキストに敏感に依存して変動すると考えられる。

以上の構成要素から、エージェント a とユーザ h の系において、エージェント a とユーザ h の感情メカニズム A, H は、以下のように記述できる。この枠組みをインタラクションも含めて、まとめると、図2のようになる。ただし、図中で \bar{e} は、 e の推定値を意味する。

なお、自分自身の表情から感情への写像 $M_{a:x \rightarrow e}^a$, $M_{h:x \rightarrow e}^h$ と、相手の感情から表情への写像 $M_{h:e \rightarrow x}^h$, $M_{a:e \rightarrow x}^a$ も考えられるが、感情マッピングの相互学習には必要ないため、枠組みには入っていない。

$$A = \langle E_a, X_a, M_{a:e \rightarrow x}, M_{h:x \rightarrow e}^a, T_a \rangle$$

$$H = \langle E_h, X_h, M_{h:e \rightarrow x}, M_{a:x \rightarrow e}^h, T_h \rangle$$

4 感情マッピングの相互学習

4.1 何が既知で、何を学習すべきか

以上のような定式化から、感情マッピングの学習 (learning of emotion mapping) とは、“相手の表情から感情への写像を学習すること”と定義される。そして、感情マッピングの相互学習 (mutual learning of emotion mapping) とは、“お互いが相手の表情から感情への写像を学習すること。つまり、エージェントは、 $M_{h:x \rightarrow e}^a$ を、ユーザは、 $M_{a:x \rightarrow e}^h$ を学習すること”である。

また、人間のユーザ h と擬人化エージェント a のインタラクションという状況では、以下のような仮定を導入できる。これらは、擬人化エージェントは少なくとも部分的には人間と同じであるという仮定に基づいている。

- (A1) 基本感情の共有: ユーザとエージェントの基本感情は同一、つまり $E = E_a = E_h$ とする。エージェントは擬人化エージェントであるので、人間と同じ基本感情が組込まれていると考えることは、妥当である。また、ユーザ、エージェントともに、自己の感情の状態は認識可能であるとする仮定も妥当であろう。

²<http://msdn.microsoft.com/workshop/imedia/agent/default.asp>

³<http://www.shadisoft.com/webspeak/index.shtml>

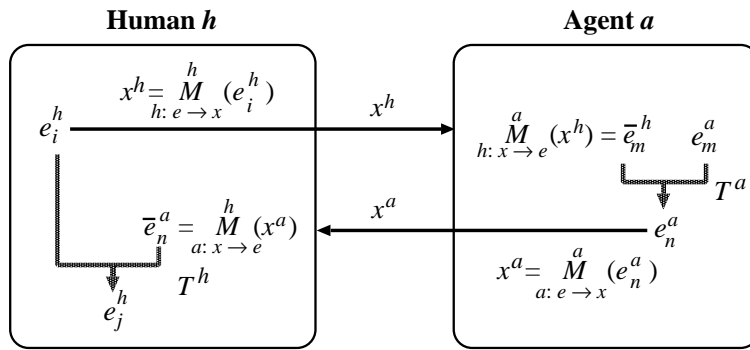


図2 エージェントとユーザのインタラクション

- (A2) 感情遷移の共有：擬人化エージェントの制約から、感情状態の遷移規則も、少なくとも部分的にエージェントとユーザが共有していると考えられる。

4.2 エージェント側の学習

ユーザ側の学習は、ユーザが任意に行うものであり、我々の枠組みとしては、それに対し何ら制約を与えない。そこで、我々が設計すべきは、エージェントの学習手続きである。

仮定 (A1) から、お互いが相手の基本感情を学習する必要はなくなる。また、相手の表情は、観測により得られるものである。よって、後は現在観測される相手の表情に対する相手の感情状態が特定できれば、単純なインスタンス学習により、エージェントは、ユーザの感情マッピングを学習することができる。

ユーザの感情状態の同定は、仮定 (A2) を基にした知識を利用する。

擬人化エージェントが実行する感情マッピングの具体的な学習手続きは、以下の方法を考えている。

1. ヒューリスティックで感情状態 e を同定して、そのときの表情 x を CCD カメラで獲得し、 $e \Rightarrow x$ をサンプリングする。
2. 集められた $\{e \Rightarrow x, \dots\}$ を事例として、 k -NN(Nearest Neighbor) 法 [2] で、 x から e を推定する。

上記のステップ 1 におけるヒューリスティックは、エージェントとユーザのインタラクション、つまりゲームに依存すると考えられるが、例えば以下のようなものが挙げられる。これらのヒューリスティックは、断片的にある程度確実な感情状態を同定するため、全体的な遷移規則があるという仮定を必要としない。以下に例を示すがおそらくこれだけでは不十分なので、さらに追加する必要があるだろう。

- 続けて (読心に) 失敗すると、“落ち込む”、“いらいらする” 状態になる。

- 調子良く進むと、“喜ぶ”。

5 相互読心ゲーム

ここでは、ユーザと擬人化エージェント間で、互いに相手の表情から感情を推定する感情マッピングの学習タスクとして、相互読心ゲームについて述べる。

5.1 相互読心ゲームの目的

相互読心ゲームの目的は、感情マッピングの相互学習のための、表情と感情のペアの教示例を効率よく、かつ広範囲に収集することである。ここで、相手の表情から感情を推定し、相手の感情を言い当てて、その正確さを競うゲームのことを、読心ゲームと呼ぶ。相互とは、エージェントが、ユーザの表情から相手の感情を推定する過程と、ユーザが、エージェントの表情からエージェントの感情を推定する過程とを、同時並行的に行うことを意味する。

一般に、こういったユーザモデルの同定のためのデータ収集過程は、エンロールと呼ばれる。エンロールの問題点は、ユーザに多数の教示データの入力を要求するため、ユーザに負荷がかかることである。そこで本研究では、ユーザの認知的負荷を軽減し、かつユーザに積極的に教示データの収集にかかわってもらうために、ゲーム形式で行なう。

一般にエージェント間のインタラクションは、競争的、協調的の 2 種に大別される。競争的読心ゲームの場合、ゲームを行う 2 者間で、どちらがより正確に、相手の感情を読んだかを競う。一方、協調的読心ゲームの場合では、ゲームを行う 2 者間が協力して、正確に、相互に相手の感情を言い当てるまでの時間の短さを他のペアと競う。どちらのタイプのゲームを感情マッピングの学習タスクとして用いる場合でも、ゲームに勝つことが、自然な感情と表情のペアの教示例の収集に影響を与えないように注意する必要がある。

5.2 協調的相互読心ゲームの概要

前述の条件を満たす擬人化エージェントとユーザのインタラクションとして、本研究では、ユーザと擬人化エージェントが協力して行う協調的相互読心ゲームを考案した。本ゲームの本来の目的は、互いの感情マッピングの同定である。そのため、ユーザ、エージェント共、それぞれの感情から表情へのマッピングを固定させてゲームを行う。特に表出する表情は、誇張されたものではなく、ユーザが普段用いる自然な表情で行う。

ゲームのルール

協調的相互読心ゲームは、以下のようなルールで進む。

1. 相手の表出した表情に対して、その時の相手の感情を推定できれば、推定した感情名を相手に対して言う。
2. 相手から感情名を言われた場合は、言い当てに対する評価として、正しい、間違っている、あるいは、近いが間違っている、の3値を、Yes/No/Near-Miss、のいずれかで、相手に伝える。
3. 上記の過程を、ゲームの終了条件を満たすまでユーザと擬人化エージェントが交互に行う。

ゲームの終了条件

擬人化エージェントとユーザのそれぞれが、連続 n 回、相手の異なる感情を言い当てれば、ゲームは終了する。言い当ての回数 n 回は、ゲームの難易度に応じて変えてよいが、最終的には、感情モデルで扱う基本感情の数とする。

この停止条件は、お互いの感情マッピングの学習が収束し、ユーザが学習した $M_{x^a \rightarrow e^a}^h$ が、エージェントの $M_{e^a \rightarrow x^a}^h$ の逆写像になり、エージェントが学習した $M_{x^h \rightarrow e^h}^a$ が、ユーザの $M_{e^h \rightarrow x^h}^a$ の逆写像になることに対応している。感情マッピングの式で記述すると以下のようなになる。

$$M_{a:x \rightarrow e}^h = \left(M_{a:e \rightarrow x}^a \right)^{-1} \quad M_{h:x \rightarrow e}^a = \left(M_{h:e \rightarrow x}^h \right)^{-1}$$

5.3 協調的相互読心ゲームの特徴

まず、ゲームの前提条件として感情から表情へのマッピングを固定させる理由は、自然な表情の教示を行うことに加えて、感情マッピングを固定させた方が、相手モデルの学習が簡単になるからである。

次に、感情マッピングの同定に、ゲーム形式を用いた意図について議論する。一般に、表情の教示データの収集において、ユーザに自然な感情を、短時間の間に、多くの種類、想起してもらうことは、容易ではな

く、作られた表情や感情になりがちである。これに対し、本研究で提案する、相互読心ゲームに基づく感情マッピングの学習では、ゲームは、終了条件を満たすまで行う。ゲーム中での、相互の言い当ての結果に応じて、言い当てが当たって嬉しい、外れて悔しい、など、ユーザ側に様々な感情が生起することを利用して、自然な感情、表情の想起を期待できる。

6 システムの実装

実際のシステムは、図3のようになる。ユーザはモニタ上に現れる擬人化エージェントをモニタし、エージェントの方は、ユーザと CCD カメラなどでモニタすることができる。後は、図中にも記述されている前述の協調的相互読心ゲームを行なうことにより、擬人化エージェントとユーザの感情マッピングの相互学習が実現される。

このシステム実装において最も技術的に問題になるのは、擬人化エージェントがもっているユーザの基本表情 X_h の設定と、ユーザの画像からの基本感情の同定であると考えられる。基本表情については、これまでの研究で用いられているものを参考にして事前に用意するか、SOM (Self-Organizing Map) [8] を用いて教師なし学習で獲得するかを検討している。また、基本感情の同定については、基本表情をどのように設定するかに依存すると考えられるが、具体的方法は今後の検討課題である。

7 まとめ

本稿では、ユーザと擬人化エージェント間でのコミュニケーションを円滑かつ柔軟にするために、相手の表情から感情を推定する感情マッピングの相互学習の枠組を提案した。そして表情から感情へのマッピングの学習のタスクとして、相互読心ゲームについて述べた。

現在システムの実装が完了しないので、今後は、本モデルを既存の擬人化エージェントに組み込み、実装し、被験者を用いた実験を行なう予定である。

参考文献

- [1] J. Bates. The role of emotion in believable agents. *Communications of the ACM*, Vol. 37, No. 7, pp. 122-125, 1994.
- [2] B. V. Dasarathy. *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*. IEEE Computer Society Press, 1991.
- [3] 土肥, 石塚. Face-to-face 型擬人化エージェント・インタフェースの構築. *情報処理学会論文誌*, Vol. 40, No. 2, pp. 547-555, 1999.

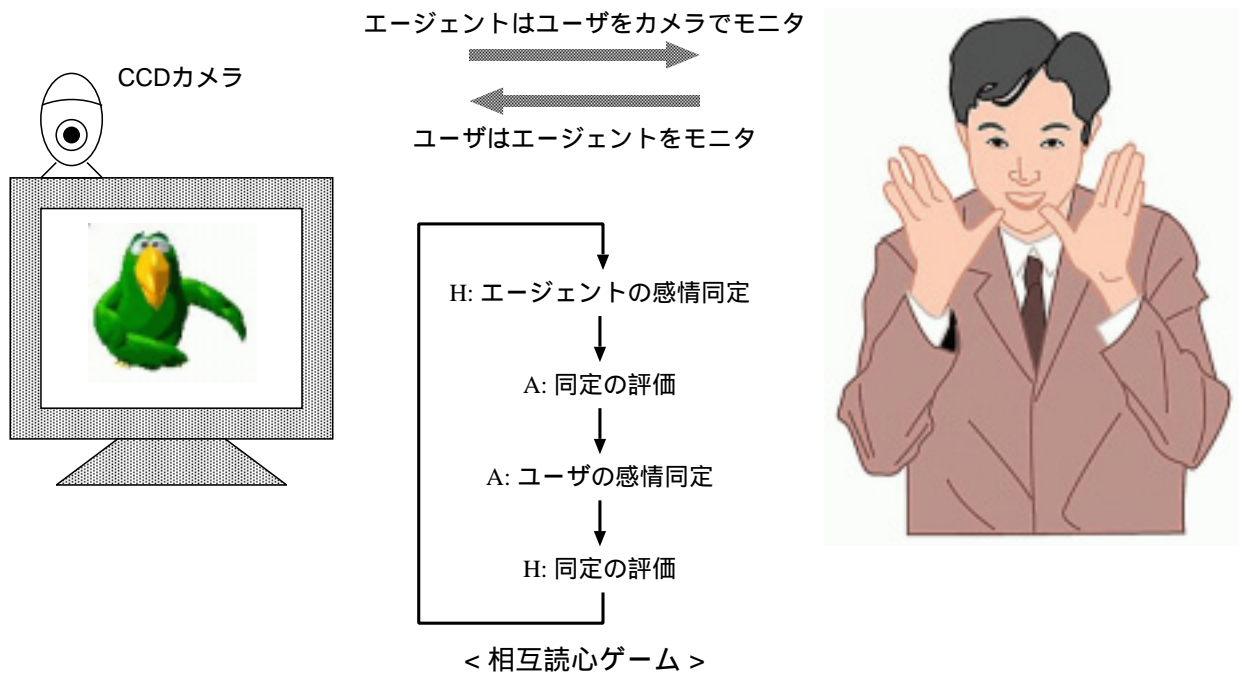


図 3 感情マッピングの相互学習システム

- [4] 飯田史也, 原文雄. 人間の教示特性に基づく顔ロボットの行動学習アルゴリズム. 日本ロボット学会誌, Vol. 18, No. 6, pp. 839–846, 2000.
- [5] 石塚満. マルチモーダル擬人化エージェントシステム. システム/制御/情報, Vol. 44, No. 3, pp. 128–135, 2000.
- [6] S. Kimura and M. Yachida. Facial expression recognition and its degree estimation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 295–300, 1997.
- [7] 小林宏, 原文雄. ニューラルネットによる基本表情認識. 計測自動制御学会論文誌, Vol. 29, No. 1, pp. 112–118, 1993.
- [8] T. Kohonen. *Self-Organization Maps*. Springer-Verlag, 1995. 徳高, 岸田, 藤村 訳: 自己組織化マップ, シュプリンガー・フェアクラーク東京 (1995).
- [9] 近藤崇, 角所考, 美濃導彦. 知的コミュニケーションのための事例に基づく表情表現モデルの作成. 電子情報通信学会信学技法, Vol. HIP99-4, pp. 23–30, 1999.
- [10] L. Steels. Emergent adaptive lexicon. In *Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pp. 562–567, 1996.
- [11] L. Steels. Perceptually grounded meaning creation. In *Proceedings of the Second International Conference on Multi Agent Systems*, pp. 338–344, 1996.
- [12] J. D. Velásquez. Modeling emotions and other motivations in synthetic agents. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, pp. 10–15, 1997.
- [13] J. D. Velásquez. When robots weep: Emotional memories and decision-making. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, pp. ???–???, 1998.
- [14] J. D. Velásquez. An emotion-based approach to robotics. In *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 235–240, 1999.
- [15] J. D. Velásquez and P. Maes. Cathexis: A computational model of emotions. In *Proceedings of the First International Conference on Autonomous Agent*, pp. 518–519, 1997.
- [16] S. Yamada and D. Kato. Acquiring a global lexicon for identifying landmarks in a heterogeneous multi-robot system. In *Proceedings of The 6Th International Conference on Intelligent Autonomous Systems*, pp. 463–470, 2000.