

感情空間を用いたロボットの表情生成とその解析

A Method to Create Expression of Robots Using Emotion Space and its Analysis

加納 政芳*1 吉田 宏徳*2 加藤 昇平 伊藤 英則
Masayoshi Kanoh Hironori Yoshida Shohei Kato Hidenori Itoh

名古屋工業大学
Nagoya Institute of Technology

The “Ifbot” robot communicates with people by considering its own “emotions” and theirs. Ifbot has many facial expressions to communicate enjoyment. We first attempted to extract characteristics of Ifbot’s facial expressions by mapping these to its emotional space, which we discuss in this paper. We applied a five-layer perceptron to the extraction. We second analyze its emotional space psychologically. We then propose a method of seamlessly changing facial expressions using the emotional space.

1. はじめに

感性会話型ロボット ifbot は、自らの感情と話者の感情を考慮に入れて会話を行うロボットである [1, 2]。ifbot は、自らの感情を表現するために数十種類の表情を持ち、これらの表情を表出しながら会話する。本稿では、ifbot の表情を感情空間にマッピングすることで、ifbot が表出する表情の特性を抽出することを試みる。特性の抽出には恒等写像学習を使用する。つぎに、得られた感情空間を解析し、感情と表情がうまくマッピングされているかを検証する。つぎに、ifbot の表情に、より違和感のない変化を持たせることを目的として、感情空間を利用したシームレスな表情表出のための手法について検討する。

2. ifbot

図 1 に ifbot の概観を示す。ifbot は、身長 45cm 重さ 7kg であり、2 つの腕を有し、車輪により移動する。ifbot には感性制御技術が実装されている。この技術は対話者の声の抑揚や調子からその人の感情を認識し、ロボットに感情を持たせる。ifbot はこの技術によって得られた感情を用いて会話を行う。

3. ifbot の表情表出機構

図 2 に ifbot の表情表出機構の概要を示す。ifbot は表情表出のために、10 のモータおよび計 102 の LED を持つ。モータ



図 1: ifbot の概観

連絡先: 加納政芳, 中京大学生命システム工学部
〒470-0393 豊田市貝津町床立 101, Tel: 0565-46-6910
mkanoh@life.chukyo-u.ac.jp

*1 現在, 中京大学生命システム工学部

*2 現在 (株)ダイアルトーンインターナショナルジャパン

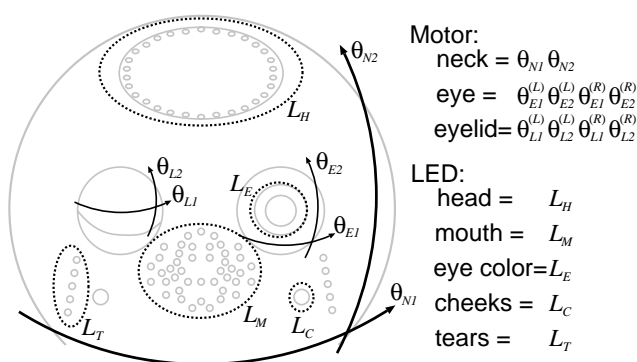


図 2: ifbot の表情表出機構

は、首を 2 軸 (同図 θ_{N1}, θ_{N2}), 左右のそれぞれの目を 2 軸 (同図 $\theta_{E1}^{(L)}, \theta_{E2}^{(L)}, \theta_{E1}^{(R)}, \theta_{E2}^{(R)}$), 左右のそれぞれのまぶたを 2 軸 (同図 $\theta_{L1}^{(L)}, \theta_{L2}^{(L)}, \theta_{L1}^{(R)}, \theta_{L2}^{(R)}$) の方向へ動作させる。LED は頭部、口部、目部、頬部、涙部および耳部の各部位に配置され、感情を表現する。LED の発色は、頭部 3 色 (橙, 緑, 赤), 口部 1 色 (橙), 目部 3 色 (緑, 赤, 青), 頬部 1 色 (赤), 涙部 1 色 (青), 耳部 1 色 (橙) である。これらの機構により, ifbot は表情豊かなコミュニケーションを行うことができる。

4. 感情空間の構築

本稿では、ifbot の表情を感情空間にマッピングするために恒等写像学習 [3] を使用する。恒等写像学習は、これまでも表情の分析や合成に用いられている [4]。恒等写像学習のネットワークは、図 3 に示すように 5 層で構成される。また、第 3 層のユニット数は入出力ユニット数より少ない構造を持つ。学習は入力データを教師信号として与えることにより行われる。学習によって、第 3 層には入力データを特徴づける情報が抽出される。本稿では、第 3 層に抽出される特徴空間を感情空間として利用する。

入力データは、10 のモータ出力値と、それぞれの部位の LED の出力値をパターン分類したデータ (図 2, L_H, L_M, L_E, L_C, L_T) とした。恒等写像学習のネットワークのユニット数は、15,45,3,45,15 とした。同ネットワークの学習には、平静時の表情 (以下, 平静) から表情変化が開始し、平静で終了する時系列データ (以下, 表情シーケンス) を使用した。

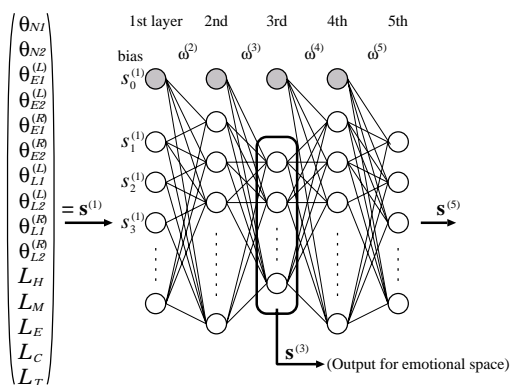


図 3: 恒等写像学習ネットワーク

表 1: アンケート結果 (%) . 各分類において, 60%以上の支持率を得た上位 2 つの表情シーケンスを示す (嫌悪, 恐れに対して 60%を越える支持率を得る表情シーケンスはなかった).

表情	怒り	嫌悪	恐れ	喜び	悲しみ	驚き	その他
exp1	78	14	0	0	4	0	4
exp2	84	4	4	2	0	2	4
exp3	0	0	0	14	0	66	20
exp4	0	0	2	22	0	72	4
exp5	2	6	4	0	86	0	2
exp6	0	0	8	2	90	0	0
exp7	0	4	0	84	0	6	6
exp8	0	0	0	96	0	4	0

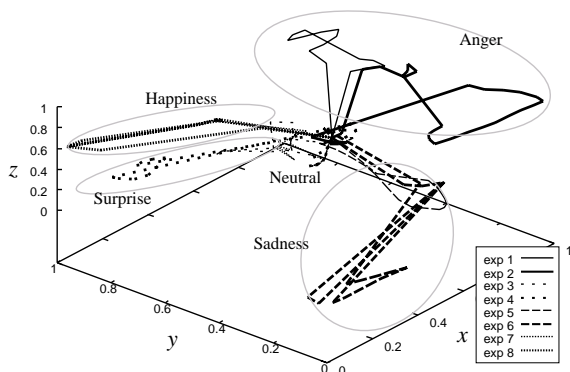


図 5: 感情空間

5. 感情空間の解析

まず, ifbot の表情シーケンスによって感情空間を解析する. これには, アンケートを用いた. アンケートでは, ifbot の表情シーケンスが提示され, 回答者はそれを基本 6 感情 [5] のいずれかの表情に類別する. アンケートは 50 人に対して行われた.

表 1 にアンケートで高い支持率を得た表情シーケンスのうちの上位 2 つを示す. 高い支持率を得た表情シーケンスの例として, 驚き (exp4) と喜び (exp8) の表情を図 4 に示す. 嫌悪, 恐れに対して 60%を越える支持率を得る表情シーケンスはなかった. 表 1 の表情シーケンスを恒等写像学習により構築された感情空間にプロットした結果を図 5 に示す. 同図より, 感情を強く表出する表情シーケンスのマッピング領域が感情空間内で分類されていることがわかる.

つぎに, 感情空間内の座標点から ifbot の表情を生成し, こ

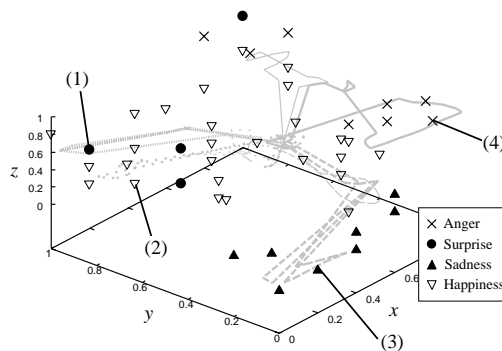


図 6: 80%以上の支持率を得た座標点

の表情が人にどのような印象をあたえるかを調べることで, 感情空間内での表情のマッピング状態を解析する. 表情の抽出は, 座標点を恒等写像学習のネットワークの第 3 層に入力し, 第 5 層から表情パラメータを出力することで行う. 計 216 の表情を感情空間から抽出し, これらが基本 6 感情のうちどれと対応がとれるかを 20 人に対してアンケートした.

図 6 に 80%以上の支持率を得た座標点を示す. 図 7 に高い支持率を得た表情を示す. ここでも, 嫌悪, 恐れに対して高い支持率を得る表情はなかった. 図 6 から, 感情を強く表出する表情シーケンスの周りには, それぞれの感情に対応する表情が分布していることがわかる. 平静付近では喜びの傾向が強く現れているが, これは ifbot の平静がほほえんだ表情 (図 7 中 neutral face 参照) を表出しているためだと考えられる. この結果から, 生成された感情空間は, ifbot の表情と感情とをうまくマッピングしているといえる.

ここで, それぞれの表情の特徴を詳しく見てみる.

怒りと識別された表情では, 口部の LED の点灯状態が図 8(A), 目部の 4 モータの出力が図 9(S) になる傾向があった. 怒りの表情は, 眉の内側が引き下がる, 口角がまっすぐ/下がるという特徴を持つ. 目部のモータと口部の LED がそれぞれの特徴を表現している. デフォルメされた口部では, 「x」の形状および強く口角が下がった形状が怒りの印象を人に与えると考えられる. また, 興奮状態を表す赤の LED が, 頭部, 目部で使用されている傾向があった. 赤色が怒りの印象を強めていると考えられる.

悲しみと識別された表情では, 口部が図 8(B), 目部が図 9(T) になる傾向があった. 眉の内側が引き上がる, 口角が下がるという悲しみの表情の特徴をこれらの形状が表現している. 口部では, 怒りの表情にみられた曲がり具合よりも小さな曲がり, 悲しみの印象を与えられる. また, 目部が青, 涙部が青 (涙を流す) に発色することで, 悲しみの感情を強く表現することができると考えられる.

喜びと識別された表情では, 口が図 8(D), 目が図 9(T) になる傾向があった. 口部の LED は, 人が喜んだ表情をすると口角が上がるという特徴を表現している. また, LED の発色状況としては頭部では緑または赤, 目部では緑が多く見られた. 悲しみと喜びの表情表出時の目部のモータ出力値には大きな差はなく, LED のみが人に与える印象を変化させていると考えられる.

驚きと識別された表情では, 口が図 8(C) になる傾向があった. 口に緊張が無く顎が落ちて開くという驚きの表情の特徴がよく現れている. それ以外の特徴は喜びと類似しており, 喜びと驚きの表情の区別は口の形状だけでされており, ifbot の表情表出機構ではその分類が難しいことがわかった.

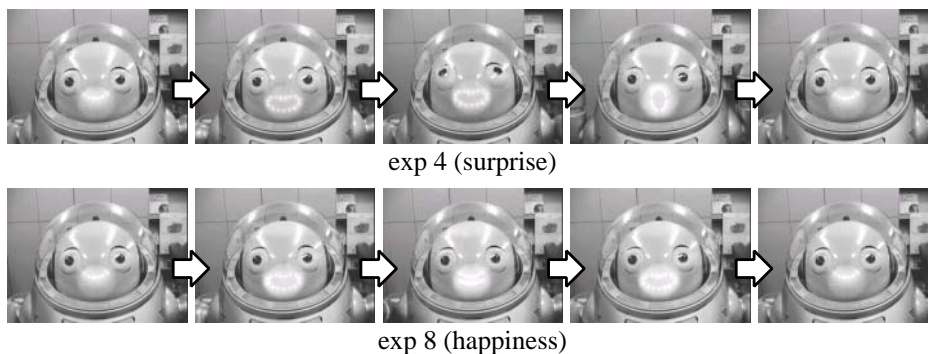


図 4: 高い支持率を得た表情シーケンス



図 7: 高い支持率を得た座標点の表情 (各画像の番号は図 6 の座標点の番号と対応) と平静の表情

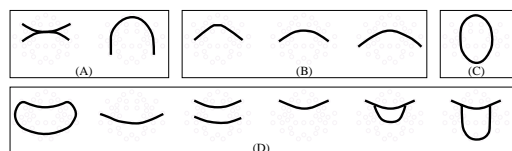


図 8: 口の形状の分類

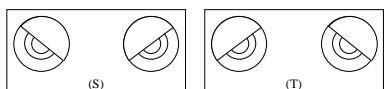


図 9: 目の形状の分類

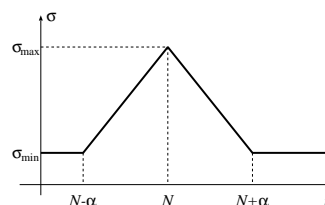


図 10: σ の動的変化

6. ガウス関数を用いた表情作成手法

アンケートで使用した表情シーケンス 2 つ連結し、表情の遷移を行わせると、中間表情として必ず平静を経由することになる。人間の表情表出の場合には、感情と感情との表情遷移の間に平静が必ず含まれるということはない。ifbot に、より自然でかつ、エンターテインメント性を持たせるためには、円滑な表情制御が必要となる。しかしながら、ifbot の表情シーケンスを直接制御することは難しい。そこで、表情シーケンスが感情空間内に描く軌跡を用いて、2 つの表情シーケンスをシームレスに連結する手法を検討する。

本稿では、シームレスな表情シーケンス作成にガウス関数による曲線平滑化を用いる。シーケンスの平滑化にはフーリエ記述子を用いることも考えられる。しかしながら、同記述子を用いた場合には、1) シーケンスの構成点数が変化する、2) 平滑化の度合いが一定であるため特徴的な表情表出が失われる、という問題点がある。これに対して、ガウス関数による平滑化ではシーケンスの構成点数は変化しない。また、ガウス関数は標準偏差 σ の値を構成点ごとに設定できるため、平滑化度を動的に変化させることが可能である。そこで、 σ を動的に変化させることで、表情の特徴を残しつつシームレスに表情を連結する手法を提案する。

具体的には、まず、感情空間内に 2 つのシーケンス $s_1(i)$ ($i = 0, \dots, N - 1$)、 $s_2(j)$ ($j = 0, \dots, M - 1$) を考える。つぎに、 $s_1(i)$ と $s_2(i)$ を連結して新しいシーケンス $s_{1-2}(i)$ を作成する。

$$s_{1-2}(i) = \begin{cases} s_1(i) & (i < N) \\ s_2(i - N) & (\text{otherwise}). \end{cases} \quad (1)$$

つぎに、得られたシーケンスの各成分 $s_{1-2}^{(j)}(i)$ (j は成分の識別子) をガウス関数で平滑化する。

$$\begin{aligned} S_{1-2}^{(j)}(i, \sigma) &= s_{1-2}^{(j)}(i) \otimes g(i, \sigma) \\ &= \int_{-\infty}^{+\infty} s_{1-2}^{(j)}(t) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(i-t)^2}{2\sigma^2}} dt. \end{aligned} \quad (2)$$

ここで、 $\sigma = \text{一定}$ とした場合には、一般的な曲線の平滑化手法となる。本稿では、 σ の値を以下の式に基づき動的に変化させる。

$$\sigma = \begin{cases} \frac{\sigma_{\max} - \sigma_{\min}}{\alpha} |i - N| + \sigma_{\max} & (|i - N| < \alpha) \\ \sigma_{\min} & (\text{otherwise}). \end{cases} \quad (3)$$

ただし、 σ_{\max} は σ の最大値、 σ_{\min} は同最小値、 α は平滑化パラメータである。図 10 に関数の形状を示す。

こうして得られたシーケンス $S_{1-2}(i, \sigma)$ から ifbot の表情シーケンスを復元したものをシームレスな表情シーケンスとする。

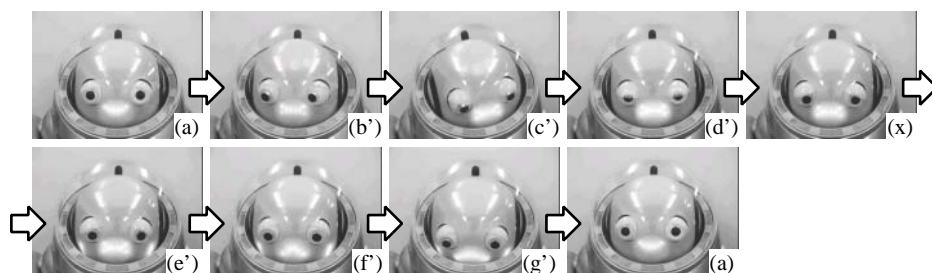


図 11: 怒り (exp2) から悲しみ (exp6) への表情変化 (提案手法)

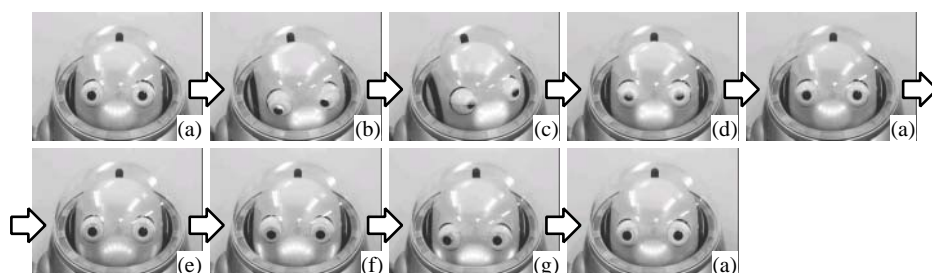


図 12: 怒り (exp2) から悲しみ (exp6) への表情変化 (単純連結). (a); 平静, (b)(c)(d); exp2 (e)(f)(g); exp6 .

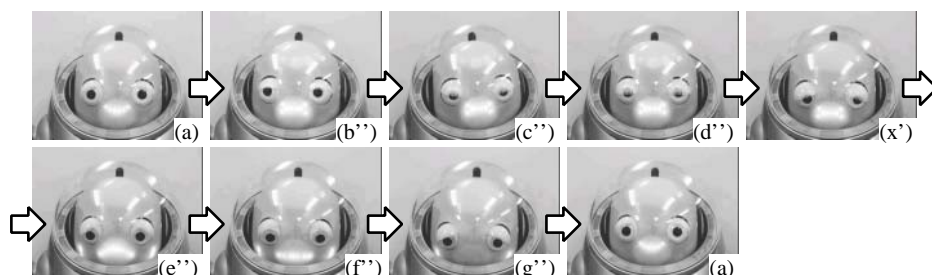


図 13: 怒り (exp2) から悲しみ (exp6) への表情変化 ($\sigma = 50$ に固定)

7. 生成された表情の評価

提案手法の有効性を検証するために表情連結の実験を行った。パラメータは $\sigma_{\max} = 40$, $\sigma_{\min} = 1$, $\sigma_{\alpha} = 150$ とした。提案手法を用いて怒りの表情 (exp2) と悲しみの表情 (exp6) を連結することで得られた表情変化を図 11 に示す。単純に 2 つのシーケンスを連結した結果を図 12 に示す。図 12 では平静 (同図 (a)) を経由しているのに対し、図 11 では同表情を経由していないことがわかる。図 13 に $\sigma = 50$ に固定したときの出力結果を示す。 σ を固定すると、特徴的な表情まで平滑化され、表情のアクティブ感を失わせていることがわかる。

8. おわりに

本稿では、恒等写像学習を用いて ifbot の表情を感情空間にマッピングすることで、ifbot の表情特性を抽出することを試みた。そして、得られた感情空間を解析し、感情と表情がうまくマッピングされていることを確認した。また、感情空間を利用したシームレスな表情表出のための手法について検討した。

謝辞

ifbot は株式会社ビジネスデザイン研究所の製品企画・総合プロデュースのもと、特に、形状デザイン、表情制御メカニズ

ムはブラザー工業株式会社、および、表情制御ソフトはブラザー工業株式会社、ロボス株式会社、名古屋工業大学が共同開発した。関連各位に感謝する。

参考文献

- [1] Ohshiro, S., Kato, S., Watabe, K., Yamakita, T., Kimura, K. and Itoh, H.: A Vision-based Talker Distinction System of Human-Robot Communication, *International Symposium on Micromechatronics and Human Science*, pp. 73-76 (2003).
- [2] Kato, S., Ohshiro, S., Itoh, H. and Kimura, K.: Development of a Communication Robot Ifbot, *The 2004 IEEE International Conference on Robotics and Automation (ICRA)* (2004). (to appear).
- [3] Bishop, C. M.: *Neural Networks for Pattern Recognition*, Oxford University Press (1995).
- [4] 坂口竜己, 山田寛, 森島繁生: 顔画像を基にした 3 次元感情モデルの構築とその評価, *電子情報通信学会論文誌*, Vol. J80-A, No. 8, pp. 1279-1284 (1997).
- [5] Ekman, P.: *Unmasking the Face*, Prentice-Hall (1975).