

平成 18 年度（2006 年度）

卒業論文要旨集

Abstracts of the 2006 Graduation Theses

岩手県立大学ソフトウェア情報学部

Faculty of Software and Information Science
Iwate Prefectural University



撮影の際には、音声と表情を的確に取得するために外付けマイクを使用し、DVでの撮影はできるだけ正面で朗読者の顔全体が入るよう設置し行った。またデータ取得時には基本6感情での分析を行うことを説明し、感情を込めた朗読の様子を撮影した。

3.2 感情のラベル付け

撮影終了後、朗読者に注文の多い料理店の文章に対して、朗読時に込めた感情を、Ekmanの基本6感情に基づくラベルを付与する形で記述してもらった。この際、同一箇所への複数の感情ラベルの記述を許可している。

3.3 INTERACT を用いた分析

3.1で取得した朗読データに対し、動画分析支援ツールであるMangold社のInteract³⁾を用いてラベル付けを行った。ラベルとしては3.2で取得した情報を元につけたEkmanの基本感情のラベル、また2.2で述べたFACSの基本単位であるAUに基づいてのラベル付けを行う。AUについては表情の動きにおいて最低限の数となる図2の基本17個のAU⁴⁾よりAU25, AU26を除いた15個を用いた。AU25, AU26は唇を開く動作となっておりラベル付け対象である朗読データにおいて、この動作は終始行われているため除外した。各AUそれぞれに対し朗読データを見ながらラベル付けを行った。その結果のを図1に示す。図1において横軸は朗読データの時間軸であり、幅の広い部分がその朗読データに対して上部より基本6感情、各AUのラベル付け個所をした部分となる。

現在ラベル付けを行った結果に対し分析を行っている。分析としては、朗読者の付けた感情ラベルに対して、FACSのAUに基づいて付けたラベルの出現傾向を比較し、賢治作品において基本6感情の表現を人がする際の特徴として抽出する事を考える。今現在これらの分析を実施している。

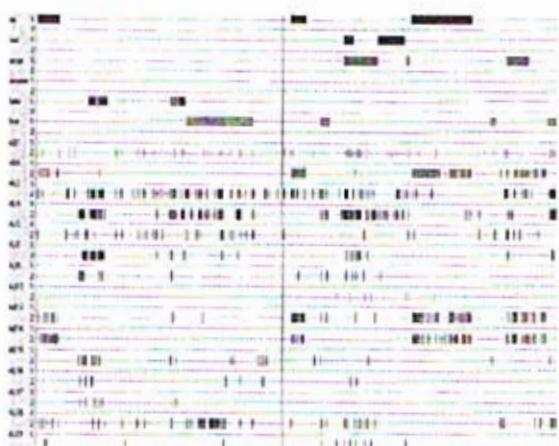


図1 基本6感情およびAUに基づくラベル付け結果

AU	AUの動作	AU	AUの動作
1	眉の内側を上げる	14	笑顎を作る
2	眉の外側を上げる	15	唇両端を引き下げる
4	眉を下げる	16	下唇を下げる
5	上瞼を上げる	17	下唇を上げる
6	頬を持ち上げる	20	唇両端を横に引く
7	瞼を緊張させる	23	唇を固く閉じる
9	鼻の皺を寄せる	25	頬を下げず唇を開く
10	上唇を上げる	26	頬を下げて唇を開く
12	唇両端を引き上げる		

図2 基本感情を表現する基本17個のAU⁴⁾

4. 終わりに

本研究では宮沢賢治の賢治らしさを表すものを賢治スタイルとし、それを構築する一手法として提案した。具体的には、賢治スタイルとして、賢治の文章に込められた感情の表現方法を捉え、複数の人間の朗読時の表情特徴を解析結果より、構築を試みる。本研究は宮沢賢治に対して実施しているが、手法が確立できれば、他の人物への応用も可能である。

今後の課題として現在実施中のデータの解析を進めるとともに、賢治スタイルの構築方法に対する検証・改善が必要であると考えられる。

謝辞

本研究を進めるにあたり、朗読データ収集にご協力いただきました宮沢賢治一軒家蔵館、また朗読風景の撮影として、賢治作品の朗読にご協力いただいた方々に心から感謝の意を表します。

参考文献

- 1) 宮沢賢治プロジェクト
<http://www.fujita.soft.iwate-pu.ac.jp/KENJI/index.html>
- 2) P. Ekman, W. V. Friesen(工藤力訳):表情分析入門 表情に隠された意味をさぐる、誠信書房(1987)
- 3) Interact -research on behavior-
Mangold International GmbH.
<http://www.mangold.de/>
- 4) 払口竜己, 山田寛, 森島繁生:顔画像を元にした3次元感情モデルの構築と評価
電子情報通信学会論文誌 A Vol. J80-A No. 8 pp. 1279-1284(1997)

専門家の朗読に基づく賢治スタイルの構築

Construction KENJI STYLE based on Acting Style Reading

インテリジェントソフトウェアシステム学講座 0312002120 成田浩規

指導教員：藤田ハミド 植松理樹

1. はじめに

現在我々は、「賢治プロジェクト」[1]と題し、宮澤賢治をコンピュータ上で再現するシステム（以後、賢治システムと記述）の構築を行っている。賢治システムの機能の一つとして、ユーザーに対し、宮澤賢治の作品を話しかけるという機能がある。ここで、朗読ではなく話しかけるという表現を使う理由は、賢治システムは、宮澤賢治自身をコンピュータ上で再現することが目的であるため、賢治システムが賢治作品を読むことは、朗読よりも話しかせるというイメージに近いためである。本機能を実現する上で重要なのは、「宮澤賢治らしさを持って作品を話しかける」ということである。「宮澤賢治らしさを持って話しかける」とは、宮澤賢治の主義思想、思考、嗜好、感情、発話の癖などに基づいて作品を表現することである。本研究では「賢治の主義思想、感情」などをまとめて「賢治スタイル」と呼ぶ。この賢治スタイルを構築するためには、賢治自身から情報を得る必要があるが、賢治はすでに他界しているため、それは困難である。そのため、現在我々が手に入れられる情報から、賢治スタイルを構築する必要がある。

以上のような背景から、本研究では、賢治に関する情報源として、宮澤賢治の専門家に着目する。宮澤賢治の専門家とは、宮澤賢治について研究している人物、賢治に対し造詣の深い人物を意味する。例えば文学者や朗読経験の多い人物がこれに当たる。本研究では、それら専門家を仮想賢治と見立てて、彼らの朗読の共通特徴を抽出することで賢治スタイルの構築を試みる。

2. 賢治スタイルの構築

2. 1 構築する賢治スタイル

本研究で構築する賢治スタイルは、文章に込められた感情から構成される。これは、賢治がその文章を書くときに持っていたと思われる感情または読者に感じてほしいと考えられる感情を意味する。

感情の定義としてはさまざまなもののが提案されているが、本研究で扱う感情は Ekman の基本 6 感情[2]とする。Ekman の基本 6 感情は、人種や地域を問わず幸福・悲しみ・怒り・驚き・嫌悪・恐怖の 6 つの基本感情に分類できることが知られている。また、他の多くの感情表出は、基本 6 感情の合成で表現できるとされている。このことから本研究においてもこの分類を用いる。

2. 2 賢治スタイル構築のための感情抽出手法

専門家（仮想賢治）の朗読について分析を行い、その結果から賢治スタイルを獲得する。今回構築する賢治スタイルには、文章に込められた感情ということから、以下に示す方法で賢治スタイルを作成する。図 1 に手法の全体像を示す。

- ① 仮想賢治による朗読の様子をデジタルビデオ（以後、DV と略記）で撮影する。
- ② 仮想賢治が文章に対して感情のラベル付けを行う。ラベルとしては、先にあげた 6 感情を用い、同じ箇所に複数のラベル付けを許す。なおラベル付けの前に朗読を実施するのは、仮想賢治のより自然な朗読を獲得するためである。
- ③ ①で得られた動画に対し、精神看護学の専門家により、感情のラベル付けを行う。
- ④ ①で得られた動画に対し、表情分析入門[2]の感情認識手法に基づき、感情のラベル付けを行う。
- ⑤ ①②の作業を複数の仮想賢治で行い、それぞれで得た情報に対し、③④の作業を行い、同じく分析する。
- ⑥ ②の感情ラベル付けを複数の仮想賢治間で比較する。複数の仮想賢治間で、同一のラベルが付けられた部分を抽出する。それらに形態素解析を行い、共通する文書特徴を取り出し、感情と対応づける。また、同一のラベルが付かなかった部分でも文章特徴を取り出し、分析する。
③④による感情のラベル付けを行う理由として、②で得られた感情ラベルは文章を対象に付けた感情ラベルであるが、実際に朗読しているときに表していた感情と、②の感情ラベルは、若干の違いがあると考えるためである。仮想賢治が朗読しているときに表していた感情を基に付けた③④の感情ラベルを作成することで、②の感情ラベルと③④の感情ラベルを比較分析し、特徴発見を試みる。

以上のように得られた情報を本研究では、賢治スタイルと認識する。

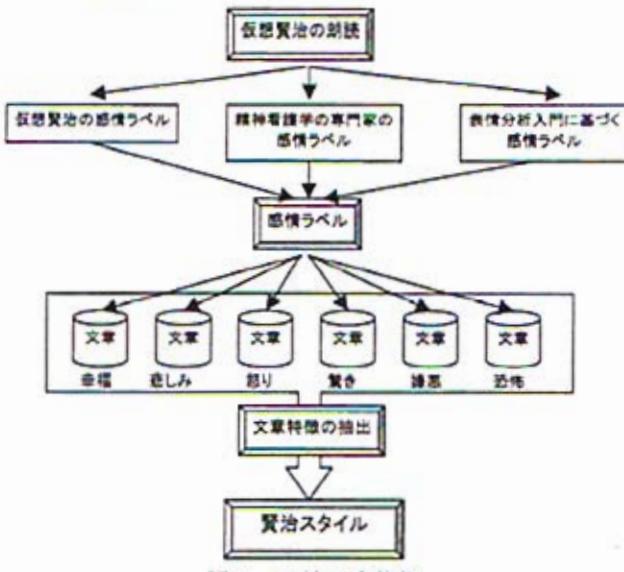


図 1 手法の全体像

3. 実験

2章で述べた手法を実現するために、仮想賢治による朗読データの収集、仮想賢治自身による感情のラベル付け、精神看護学の専門家による感情のラベル付け、表情分析入門に基づく感情のラベル付けを行った。

今回の収集では、仮想賢治として、宮澤賢治の研究を行っている文学者・佐々木民夫氏（岩手県立大学社会福祉学部・教授）、宮澤賢治についての造詣の深い牛崎敏哉氏（宮澤賢治イーハトーブ館・副主幹・学芸員）宮澤賢治の研究を行っている文学者・望月善次氏（岩手大学教育学部・教授）に協力をお願いした。特に牛崎氏は、賢治作品の朗読に定評のある人物である。また、仮想賢治が朗読している様子を撮影した動画を見て、感情のラベル付けを行う精神看護学の専門家として、樋口日出子氏（岩手県立大学看護学部・助教授）、伊関敏男氏（岩手県立大学看護学部・助手）に協力をお願いした。

3. 1 朗読の撮影

朗読する作品として、「注文の多い料理店」[3]を選んだ。選定理由としては、賢治の生前に出版されている数少ない作品の一つであること、物語を通して感情の移り変わりが多くあるということがあげられる。撮影の際には、音声と表情を的確に採取するために外付けのマイクを使用し、DVでの撮影対象は仮想賢治の顔全体が入るようにして、仮想賢治にはなるべくDVの方を見ながら朗読を行うように指示した。また仮想賢治には基本6感情に基づき分析を行うということを説明し、それを意識した上で朗読を行った。

3. 2 仮想賢治自身による感情のラベル付け

朗読終了後、仮想賢治に注文の多い料理店の文章に対し、朗読時に込めた感情を、Ekmanの6感情に基づく感情ラベルを付与することで標記した。この際、同一箇所に複数の感情ラベルを付与することを許すとともに、できるだけすべての文章にアンダーラインを示すように指示した。

3. 3 INTERACT[4]を用いた感情のラベル付け

3. 1で得た動画を基に、樋口氏と伊関氏に2人で1つの感情ラベルを付けた。その際、動画分析支援ツールであるMangold社のINTERACTを用いて感情のラベル付けを行った。図3にINTERACTを用いたラベル付け結果を示す。なお実際の画面には図3下側にある文章は表示されない。

また、表情分析入門に基づく感情のラベル付けに対しても、INTERACTを用いて感情のラベル付けを行った。

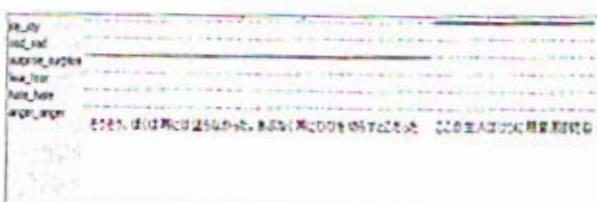


図2 INTERACTでのラベル付け（拡大）

3. 4 分析

今回、得られている感情ラベルは、牛崎氏による感情ラベル(U1)、牛崎氏の朗読の様子を撮影した動画を基に付けた、精神看護学の専門家による感情ラベル(U2)、また表情分析入門に基づく感情ラベル(U3)、佐々木氏による感情ラベル、望月氏による感情ラベルである。

現在、分析では、U1,U2,U3の比較分析を行っている。まず、注文の多い料理店のすべての文章に対し形態素解析を行い、さらに、U1とU2とU3、U1とU2、U1とU3、での共通ラベル部分の文章を抽出し、それぞれを形態素解析した文章をごとに分類した。その際、形態素解析ツールである奈良先端技術大学院大学の茶筌[5]を使用した。

4. 今後の課題

現在、佐々木氏、望月氏の朗読の様子を撮影した動画を基に、精神看護学の専門家による感情のラベル付け、表情分析入門に基づく感情のラベル付けを行う。それにより、仮想賢治による感情のラベル付けと合わせて分析し、同一のラベルが付けられた部分を抽出し、そこに含まれる語句を中心に共通特徴の発見、また同一のラベルが付かなかった部分でも、同じく分析を試みる。さらに仮想賢治ごとに同じ感情ラベルがつけられた範囲を比較することも実施する。現在これらの分析を実施している。

5. おわりに

本研究では、宮澤賢治の賢治らしさを表すものを賢治スタイルとして捉え、それを構築する手法について提案した。具体的には、賢治スタイルとして、文章に込められた感情を賢治の専門家の朗読を解析した結果に基づき、構築を試みる。本研究は宮澤賢治に対し実施しているが、手法が確立できれば、他の人物への応用も可能である。今後の課題としては、現在実施中のデータの収集・解析を進めるとともに、賢治スタイルの構築方法の検証・改善が挙げられる。

謝辞

本研究を進めるにあたり、朗読データ収集への協力、感情のラベル付けへの協力、専門家からの多くの有益なコメントをいただきました。岩手県立大学社会福祉学部教授・佐々木民夫氏、宮澤賢治イーハトーブ館副主幹・学芸員・牛崎敏哉氏、岩手大学教育学部教授・望月善次氏、岩手県立大学看護学部助教授・樋口日出子氏、岩手県立大学看護学部助手・伊間敏男氏に心から感謝の意を表します。

参考文献

- [1] 宮澤賢治プロジェクト, <http://www.fujita.soft.iwate-pu.ac.jp/KENJI/index.html>
- [2] P.Ekman,W.V.Friesen(工藤力訳), 表情分析入門, 誠信書房, 1987.
- [3] 宮澤賢治, 「注文の多い料理店」, 岩崎書店, 2004
- [4] 分析ツール, INTERACT, <http://www.mangold.de/>
- [5] 形態素解析ツール, 茶筌, <http://chasen.naist.jp/hiki/Chasen/>

動画像を用いた表情からの感情推定手法

Inference of Emotion on Facial Expression from Image Sequences

インテリジェントソフトウェアシステム学講座 0312003031 柏倉守
指導教員：藤田ハミド 羽倉淳

1. はじめに

岩手県立大学ソフトウェア情報学部藤田研究室では「宮澤賢治プロジェクト」¹⁾と題してユーザーとインタラクティブな対応をする仮想宮澤賢治の作成に取り組んでいる。本研究は宮澤賢治プロジェクトにおいて表情認識部にあたる。

HCI(Human Computer Interaction)への応用を目的として、表情から感情を認識するシステムの研究が行われている²⁾³⁾。従来の手法では FACS(Facial Action Coding System)⁴⁾で定義されている基本 6 感情(驚き、恐怖、嫌悪、怒り、幸福、悲しみ)のうちいずれか一つの感情認識に留まっていることが多い⁵⁾。しかし人の表情は単一の感情を示していることだけではなく、複数の感情が混ざった表情をしていることも多く見られる。以下、本研究では複数の感情が混ざった感情を混合感情と呼ぶ。

そこで本研究ではユーザーの表情を眉・目・口の 3 パーツにわけて各パートが表す感情を認識し、それらの感情をもとに単一の感情または混合感情、意図を推定する。各パートで認識する感情は基本 6 感情と中立とする。2.2 節で述べるように各パートが表す感情の組み合わせによっては各パートに表れた感情と顔全体として表現された感情が異なる場合もある。そのため本研究ではパート毎に表情認識のテンプレートを導入し、パート毎に表出された感情を認識する。そのうえで認識されたパート毎の感情の組み合わせからルールにより顔全体が表現する感情を推定する。

2. 感情の推定手法

本研究ではパート毎の感情認識と各パートが表す感情の組み合わせから顔全体で表す感情の推定の 2 つに大別できる。システムの全体像を図 1 に示す。

2.1 パート毎の感情認識手法

パート毎で表される感情を認識するためのテンプレートの作成とテンプレートを用いたパート毎の感情認識を行う。

2.1.1 表情からの感情認識テンプレート

本研究でのテンプレートとは基本 6 感情のいずれかのラベルとパートの動き方を予測する線形システムから成る。本研究では表情譜⁶⁾のようにパート毎のある動き方を 1 つの線形システムとみなし、その線形システムによって感情を表すパートの動きを定義する。線形システムは遷移行列を持ち、遷移行列により顔の特徴点の座標から 1 フレーム後の座標を予測できる。ここで顔の特徴点は各パートの動きを捉えるために定めたものであり、本研究における特徴点の数は左右の口で 6 点、左右の目で 8 点、口を 8 点とした。ま

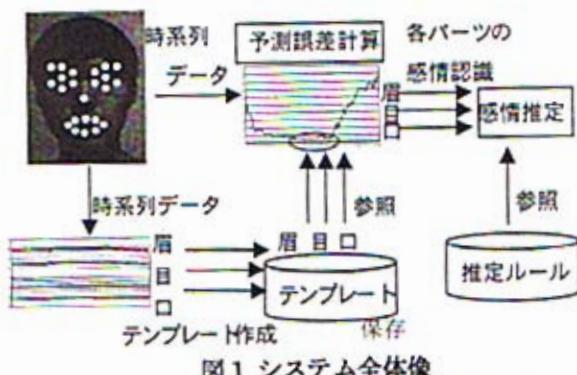


図 1 システム全体像

た感情を示す表情変化と無関係な顔の動きを除くための基準点として、表情変化による動きが少ない鼻にも 1 点設ける。特徴点の位置を図 1 の顔写真に示す。テンプレートの作成手法はまず表情の時系列データからテンプレートにする区間を指定し遷移行列を作る。また指定されていない区間も同様に遷移行列を作る。各区間の遷移行列の類似度に基づき、指定区間と指定外の区間が同定されないようにクラスタリングを行う。それにより作られた指定区間の遷移行列をテンプレートに用いる。

2.1.2 テンプレートマッチング手法

各パートの表情の動きから感情を認識するために認識する表情とテンプレートのマッチングを行う。

マッチングの手法はまずテンプレートが持つ線形システムにより時系列データの 1 フレーム後の特徴点の座標を予測する。次に予測した値と実際の値との誤差(予測誤差)を求めるで行う⁷⁾。

予測誤差が小さければ、認識している表情の動き方とテンプレートで定義された動き方が似ているということである。よって、予測誤差の小さいところをそのテンプレートにラベル付けされた感情とする。閾値を与え、複数のテンプレートのうち予測誤差が最小のものが閾値よりも小さければそのテンプレートのマッチング度合いが一番高いとし、そのテンプレートにラベル付けされた感情と認識する。基本 6 感情が認識されなかったときの感情は中立とする。

2.2 各パートで認識された感情からの顔全体の感情推定

各パートで認識された感情をもとに、認識された顔全体がどのような感情・意図を表しているのかを推定する。ここで意図とは、基本 6 感情とは異なる疑問といったような付加的情報のことと指す。推定は推定ルールに基づいて行う。推定ルールは心理学に基づき、各パートから認識された感情の組み合わせとその推定結果となる感情と意図を定義しておく。顔全体からの感情推定には(1)各パートで認識された

感情の和から求める。しかし顔全体で見たときは(2)各パーツに表れている感情を感じていない場合もある。また、(3)意図を示す表情もある。そのため推定ルールとして(2)、(3)のときの組み合わせのパターンとその推定結果となる感情・意図を決める。それに当てはまらないときは(1)とし各パーツで認識された感情の和を推定結果とする。(1)、(2)のときは感情、(3)のときは意図または感情と意図が表示される。

3. 評価実験

テンプレートの作成とテンプレートを用いたマッチングを行った。テンプレートは1人の実験参加者に基本6表情を演技してもらい、そのデータから作成した。また同じ人による怒りの演技表情とのマッチングをした。データの取得にはモーションキャプチャシステム(EVa HiRES Version 7.0)⁸⁾を使用した。

3.1 結果

作成されたテンプレートの示す動きとテンプレートが作成された実際のデータの動きを比較したものを見ると図2、図3に示す。またマッチングの結果として、怒りの目に対する6感情分のマッチングを行った結果を図4に示す。この結果では表情変化部において怒りのテンプレートの予測誤差が最も低い値をとっていた。しかし表情変化部の始めのほうでは他の感情のテンプレートも同等の予測誤差であり、怒りの予測誤差よりも値が小さいものもあった。

3.2 察察

テンプレートでの動きは実際の動きより直線的な軌跡になっている。これはテンプレートでは表情動作を線形システムとして扱っているためである。よって実際の動作が直線的なほどテンプレートは動作を再現することができる。

マッチングの結果より認識を行う際には閾値をえて閾値以下の最も小さいものを認識するのだけではなく、前後の予測誤差を用いて認識を行うことにより正確な認識が可能になると考えられる。またデータの補正を鼻の基準点によるものだけではなく正確な補正をすることで認識率が上がると考えられる。

4. おわりに

本研究では顔のパーツ毎の感情認識と、顔全体の表情からの感情の推定手法の提案・構築・推定ルールの作成を行った。表情からの感情を認識するためにマッチングを行ったが、認識精度が低かった。そのため各パーツから顔全体の感情を推定するには至らなかった。

今後はパーツごとの認識精度を高くするために時系列データの座標補正を行う。また、テンプレートの作成手法、テンプレートとのマッチング手法の提案・検討も継続して行う。また、宮沢賢治プロジェクトに応用するためには自然な表情の認識、6感情に留まらない感情の推定を個人差による影響を受けずに行う必要がある。今後もそれらを考慮してシステム構築を行っていく。

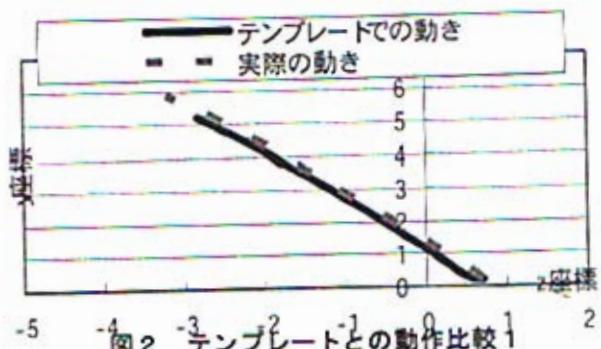


図2 テンプレートとの動作比較

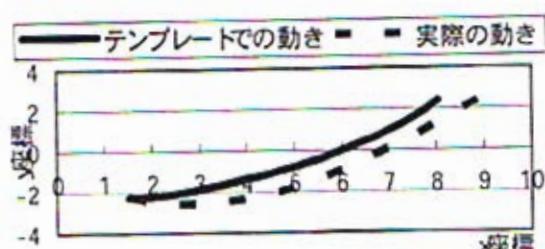


図3 テンプレートとの動作比較

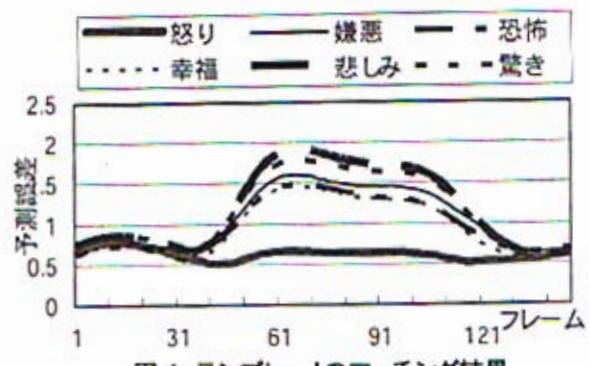


図4 テンプレートのマッチング結果

参考文献

- Hamido Fujita, 他 2 名: Virtual Cognitive Model for Miyazawa Kenji Based on Speech and Facial Images Recognition, WSEAS Transactions on Circuits and Systems, Issue 10, Vol.5, pp.1536-1543(2006).
- 下田宏, 他 2 名: 動的顔画像からのリアルタイム表情認識システムの試作, ヒューマンインタフェース学会論文誌 Vol.1, No2, pp.25-32 (1999).
- 叶冠峰, 他 2 名: 顔動画像からの特徴点抽出を用いた表情認識, 岐阜大学大学院工学研究科 (2005)
- Paul Ekman, Wallace V. Friesen, 工藤力 訳: 表情分析入門, 誠信書房 (1987).
- Maja Pantic, 他 : Automatic Analysis of Facial Expressions: The State of the Art, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL.22 NO.12(2000).
- 西山正紘, 他 2 名: 表情譜・顔バーツのタイミング構造の記述とその自動獲得, 京都大学大学院情報学研究科 (2005).
- 山平慎吾, 他 2 名: 表情譜を用いた自然な表情の認識手法, 岩手県立大学ソフトウェア情報学部 (2006).
- MotionAnalysis Corporation: EVa 7.0 Reference Manual, 住商エレクトロニクス (株), (2002).

交流分析を用いたインタラクションモデルの構築

An Interaction Model based on Transactional Analysis

インテリジェントソフトウェアシステム学講座 0312003067 佐々木瞬
指導教員：藤田ハミド 羽倉淳

1. はじめに

人間とコンピュータが人間同士と同等のインタラクションを行うためには、人間がコンピュータに対して感情移入できる事が不可欠な要件の一つである。そのためにはコンピュータがあたかも個性を持つように振舞う必要がある。現在、我々が実施している「宮沢賢治プロジェクト」¹⁾はそのようなシステムの一つである。すなわち、ユーザの表情や音声などの様々な行動からユーザの心理状態を推定し、その状態に合わせて読み方や物語などを変えて朗読することを目的とする。

インタラクションシステムは様々な場所で研究開発されているが、それらのシステムは入力に対する出力の組み合わせが決められており²⁾、ユーザとのやり取りの中で入出力の組み合わせが変化することがない。

そこで、本研究では入出力の組み合わせが変化するように、システムの内部状態が出力に影響を与えるインタラクションのモデルを考える。ここでは交流分析の知見を用いて、対話における内部状態の組み合わせを利用する。この組み合わせをベースにし、賢治本人の個性を反映させた対話パターンに基づくインタラクションを行うモデルの構築を試みる。

2. インタラクションモデル構築

2. 1 交流分析

はじめに本研究においてインタラクションのベースとして用いた交流分析を紹介する。交流分析ではすべての人は3つの自我状態を持つとされている³⁾。それらはP(Parent)：親の自我状態、A(Adult)：大人の自我状態、C(Child)：子供の自我状態と呼ぶ。またPとCの自我状態はそれぞれCP(Critical Parent)：批判的な親とNP(Nurturing Parent)：保護的な親、FC(Free Child)：自由な子供とAC(Adapted Child)：順応した子供の2つに分けられる。交流分析では人の性質や表情や声などの特徴をこれらの5つの自我状態に分類している。

交流分析ではこの5つの自我状態を使ってエゴグラムを作る。エゴグラムは人の個性や特徴などを簡単なテストにより自我状態のエネルギー量としてグラフ化する手法であり、これを利用することで、自己変

容を行いややすくする。

また、交流分析では対話の中で人がどのように作用するかを分析する交流パターン分析という手法がある。交流パターン分析では、あらゆる時点において個人が持つ自我状態のうち活動しているのは一つだけであるとされ、2者の対話における自我状態の全ての組み合わせの中で、特定のパターンにより対話の継続しやすさが分かるとしている。また、特定の自我状態からの対話に対する相手の受け方に、最も起こりやすい組み合わせが存在することも示している。

2. 2 インタラクションモデルの構成

本研究で構築するインタラクションモデルは、ユーザの自我状態と感情、システムの自我状態と感情、システムが個性として持つエゴグラム、賢治スタイルを要素として持つ。ここでいう賢治スタイルとは、宮沢賢治作品の文に対して基本6感情の何れかが関連付けされており、作品を朗読する際にその感情を込めることで賢治らしい朗読が表現できるのもとして定義してある。ユーザとシステムの感情はEkmanの基本6感情⁴⁾によって定義する。基本6感情は人種や地域を

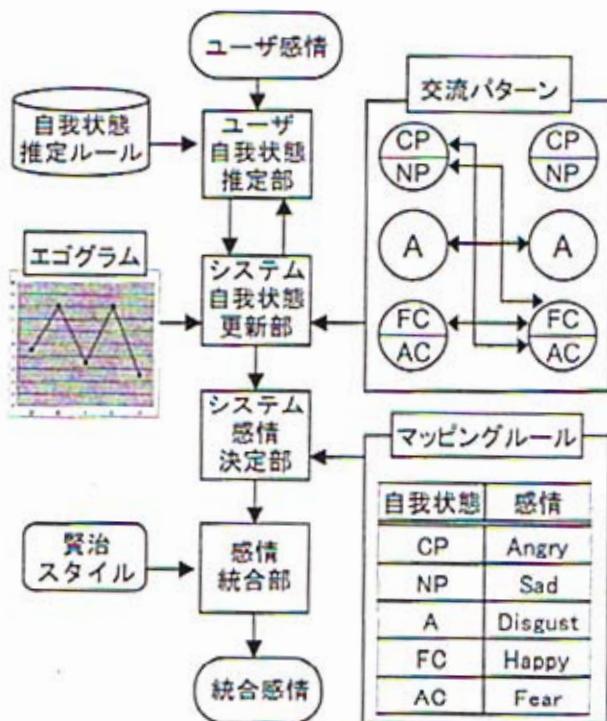


図 1 インタラクションモデルの処理フロー

問わず 6 つの基本的な感情に分類できることが知られている。これに加え、交流分析の知見を用いたルールと関数を使ったインタラクションモデルの処理フローを図 1 に示す。

2.3 インタラクション方法

図 1 で示したモデルの詳細を述べる。ここではシステムの自我状態を S_{in}^{system} 、システムの感情を E_{in}^{system} 、ユーザーの自我状態を S_{in}^{user} 、ユーザーの感情を E_{in}^{user} 、システムの持つエゴグラムを ego 、賢治スタイルが示す感情を E_{in}^{kenji} 、システムの出力感情を E_{in}^{output} とする。 t_0 は t_0 から始まる時系列のラベルである。

本研究で構築するインタラクションモデルは入力としてユーザーの感情を受け取り、システムの感情と賢治スタイルで示される感情の統合したものを出力する。以下に各処理部の詳細を述べる。なお、本手法ではシステムとユーザーの自我状態の判別に互いの自我状態が関係していることから、初期処理としてシステムの自我状態 $S_{t_0}^{system}$ を任意の 1 つに決定している。

ユーザー自我状態推定部：

①式で示すとおり E_{in}^{user} と S_{in}^{system} から S_{in}^{user} を求める。 $F(\cdot)$ は自我状態推定ルールである。交流パターン分析で示される 2 者の自我状態間の特徴から受け手の感情として最適だと思われるものを定義し、それらをルールとして作成する。このルールに従い S_{in}^{user} を求める。

$$S_{in}^{user} = F(S_{in}^{system}, E_{in}^{user}) \quad ①$$

システム自我状態更新部：

②式で示すとおり S_{in}^{system} と S_{in}^{user} と ego から $S_{t(n+1)}^{system}$ を求める。 $G(\cdot)$ は自我状態更新ルールである。更新は 3 手順で行われる。i) 交流パターン分析で示されている最も自然に起こりやすい組み合わせを一般的な対話ルールとし、対話ルールにおいて S_{in}^{user} の対になる自我状態を S_{in}^{system} とする。ii) ego の各自我状態の値の算術平均を Ave とし、 ego の S_{in}^{system} に該当する自我状態の値に Ave を加え S_{in}^{system} に該当する自我状態の値に $Ave/2$ を加える。iii) 変化した自我状態の値の最大値を示す自我状態を $S_{t(n+1)}^{system}$ とする。

$$S_{t(n+1)}^{system} = G(S_{in}^{system}, S_{in}^{user}, ego) \quad ②$$

システム感情決定部：

③式で示すとおり S_{in}^{system} から E_{in}^{system} を求める。 $H(\cdot)$ は自我状態と感情のマッピングルールである。交流分析で示されている 5 つの自我状態の表情や言葉、性質などからそれらに最も適した感情を

規定し、自我状態と感情をマッピングする。このルールにより E_{in}^{system} を求める。

$$E_{in}^{system} = H(S_{in}^{system}) \quad ③$$

感情統合部：

④式で示すとおり E_{in}^{system} と E_{in}^{kenji} から E_{in}^{output} を求める。統合ルールの基本的な考え方「システムは E_{in}^{kenji} を積極的に表現する」である。 W_{in}^{kenji} は E_{in}^{kenji} の強度を示し $0 < W_{in}^{kenji} \leq 1$ の値をとる。

④式は E_{in}^{kenji} と W_{in}^{kenji} によって強度が変化した E_{in}^{system} を統合させるという意味であり、+ は感情の統合を、- は減算を示す。

$$E_{in}^{output} = E_{in}^{kenji} + W_{in}^{kenji} (1 - W_{in}^{kenji}) E_{in}^{system} \quad ④$$

3. 評価実験

本モデルは各個人の特徴を反映したインタラクションを実現するものである。これを確認するためにモデルの対象となった個人とモデルの反応を比較する。本実験では、提案手法の基本性能を確認するために、実在する被験者 5 名を対象とし、その比較結果により、本モデルの有効性を検証した。詳細は本論に示す。

4. 終わりに

本稿では交流分析における自我状態による交流パターンとエゴグラムを利用したインタラクションのモデルを提案した。本手法はシステムの自我状態とユーザーの自我状態の関係性を基にインタラクションを行なうことを試みるものである。

今後の課題は、感情の表現方法を変えることである。具体的には同時に複数の感情が複合した形として持つようにすることや 6 感情ではなくより多くの感情表現を使えるようにすることである。これにより、各自我状態と感情の関係性が明確になり、より自然なインタラクションを起こすことが出来ると考える。

参考文献

- 1) 宮沢 賢治 プロジェクト：
<http://www.fujita.soft.iwate-pu.ac.jp/KENJI/index.html>
- 2) 中沢 正幸、西木 卓也、嵯峨山 茂樹：視線制御モデルによる擬人化エージェントの制御、人工知能学会 3B2-07 2005.
- 3) 澤田 一哉他：ヒューマノイドエージェントの構成技術と応用に関する研究、戦略的情報通信研究開発促進制度 (2004)
- 4) 杉田峰康：交流分析、日本文化科学社 (1985)
- 5) P. Ekman 他、工藤力 訳：表情分析入門、日本文化科学社 (1985)

表情特徴に基づく感情推定に関する研究

Emotion Recognition Based on Features of Facial Expression

インテリジェントソフトウェアシステム学講座 0312003080 十文字俊裕
指導教員：羽倉淳 橋松理樹 藤田ハミド

1. はじめに

岩手県立大学ソフトウェア情報学部藤田研究室では「宮沢賢治プロジェクト」¹⁾を進めている。本プロジェクトでは、仮想の宮沢賢治を実現し、人との対話などのコミュニケーションを行うことを目的としている。人同士のコミュニケーションにおいても重要な役割をする表情に注目し、表情認識からの感情推定を行う。

表情を認識する際の特徴は、顔の表情筋によって動く眉・目・口の動作や皺によって表情が作られている。表情認識の代表的手法として、Ekmanらによって提唱された Facial Action Coding System (FACS)²⁾で記述されている Action Unit(AU)と呼ばれる視認できる最小の表情動作を認識し、AUの組み合わせから基本 6 表情(幸福、驚き、恐怖、怒り、嫌悪、悲しみ)³⁾を認識する手法が広く利用されている。従来の研究でも、AUを用いた表情認識するシステム⁴⁾が作成されているが、まだ、AUや感情を混同して認識することが問題点としてあげられる。また、対話などのコミュニケーションをしている際の感情推定をする場合には、話している口の動作から感情を認識することは難しく、顔の上部である眉、目部分の AU の判断が重要となる。

そこで本研究では、ユーザの表情から従来の方法である AU の認識に加え、各表情の特徴として挙げられている、AU にはない額の皺や眉間の皺⁵⁾からも、感情を推定することで、感情を明確に推定することを試みる。

2. 感情推定手法

本研究では、グレイスケールの動画像から表情の動作と皺を認識し、感情推定をするシステムを構築する。本手法では、特徴点認識、表情動作認識、皺認識、感情推定の 4 つ部分からなるシステムを構築する。システム全体像を図 1 に示す。以下に各部分の説明を述べる。

2. 1. 特徴位置認識

ここでは、カメラから得られた動画像から AU を認識するための特徴点や皺の出る領域を認識するために、Active Appearance Model (AAM)⁶⁾を使用し、動画像からの特徴点位置の認識を行う。AAM では、AU を認識するための特徴点を設定した学習サンプルに最も近い特徴点の位置を見つけることができ、さらに、その特徴点の追跡を行うことができる。特徴点の位置は眉に 3 点、目に 4 点、口に 4 点と輪郭線に 13 点の 31箇所に設定する。表情動作認識には、目、眉、口の 18 点から行う。また、皺を検出する領域を認識された特徴点の位置から判断する。表情動

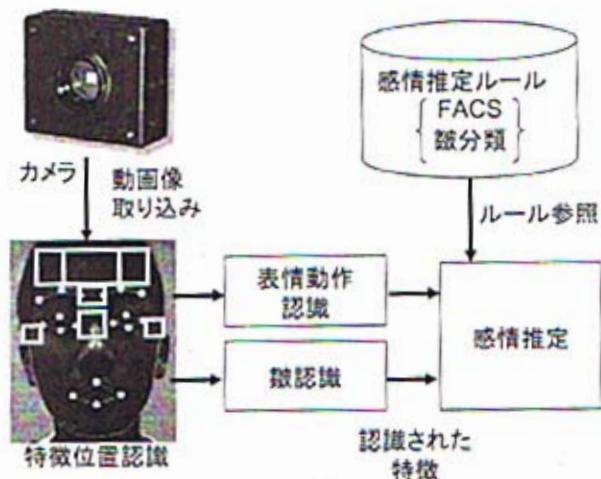


図 1 システム全体像

作認識に利用する特徴点の位置、皺認識位置を図 1 内に示す。

2. 2. 表情動作認識

ここでは、AAM によって認識された、特徴点の位置から、表情の動作をから AU の認識を行う。AAM によって認識された特徴点から、その特徴点の中で、表情変化によってほとんど動かない特徴点を結んだ線を基準線として設定し、そこからの目・眉・口の各特徴点との距離から特徴量を求める。基準線は、(1)両目の目尻を結んだ線(2)片目の両端を結んだ線(3)口の両端を結んだ線を設定し、(1)からは、眉と口の両端の動作(2)からは目の開閉の動作(3)からは口の開閉の動作の判断をそれぞれ行う。特徴量の各値には、中立表情の複数の顔画像の各特徴点の位置からニュートラルの値を算出し、その値からの増加・減少により眉の内側を下げるなどの表情動作の判断をする。また、複数の顔画像から特徴量の最大値・最小値も求め、その割合から AU の動作の大きさを求める。特徴量から基本 6 表情に関係のある 18 個の AU 動作パラメータを決定する。AU 動作パラメータは、各 AU の大きさを表す。

2. 3. 皺認識

ここでは、基本 6 表情に関係する皺の認識を行う。皺は、AU に含まれる皺と AU には含まれないが表情と関係している皺の 2 種類を認識する。表情に表れる皺は、Ekman らが行った表情分析で表情ごとの特徴として上げられている皺を基に感情ごとに分類する。驚き・怒り・幸福・嫌悪・恐怖の感情を判断するために、額中央、額両端、眉間、鼻の付け根、目尻の 7 箇所から皺の成分を抽出する。悲しみからは

感情	出現位置
驚き	額両端・中央
怒り	眉間
恐怖	額中央
幸福	目尻
嫌悪	鼻

表1 感情ごとの皺領域の分類

日立った皺を認識しづらいため、基本6表情から悲しみを除く5つの感情を判断する。認識する各表情の皺領域の分類を表2に示す。

認識された顔の特徴点の位置から皺の出る領域を判断し、その領域に対してテクスチャの画像に、Gaborフィルタを適用し、皺成分の抽出を行う。Gaborフィルタは、特定の方位の輪郭線を抽出する画像処理で、皺が額を横切る皺や眉間の縦皺など、一定の方向の成分を抽出し、皺を認識しやすくするために利用する。各皺の領域で、フィルタを適用し、抽出された方位の成分から一定量の連続量として成分が出た場合にそれを皺として認識する。各皺領域に、皺パラメータを定義する。皺パラメータは皺のon/offにより定義する。

2.4. 感情推定

ここでは、表情ごとの動作や皺出現位置のパターンから感情推定のルールを作成し、入力されたパラメータを照合して感情を認識する。感情推定のルールはFACSに基づいて基本6表情ごとの各AU動作パラメータのパターンからルールを作成する。感情ごとの皺出現位置のパターンからもルールを作成する。

推定手法は、表情動作を目・眉の動作を含む上部と口の動作の下部に分る。皺からも感情を推定しその3領域からそれぞれ表情ごとのルールにマッチした感情を推定する。

出力する感情は、動作から抽出された感情に対し皺がどのような影響があるか検証し、顔全体の、表情上部・表情下部・皺の感情ごとの占める割合をそれぞれ定義し、複数の感情を出力できるようにする。検証の際には表情動作、皺それぞれ感情を出力し、出力された感情を分析する。

3. 実験

表情を変化させた動画像を撮影し、2章で述べた手法でシステムを実装し、本手法の有用性を図る評価実験を行う。実装したシステムを用いることで、カメラから撮影された動画像から感情が表出している区間において表情動作、皺からの感情推定を行う。また、表情動作のみの場合と皺もあわせた感情推定を比較し、皺の有効性も検証する。実験は、実験参加者の前にCCDカメラ(Point Grey Research Inc.製Dragonfly Express)を配置し、顔の正面からの動画像を撮影し、動画像の各フレームに対して感情の認識を行う。動画像のレートは毎秒30フレームで撮影す

る。実験参加者には、大学生を想定し各表情の動画像を撮影する。撮影した動画像に対しては、認識した感情と表情動作を出力する。

4. 今後の展開

現在、撮影した動画像(AVI)から抽出した各フレームのピットマップ画像に対し、AAMを利用して特徴点を認識するプログラムを作成した。出力される点の座標位置から、特徴量を算出し感情を推定するプログラムを作成し、実験を行う。今後は、識別部の作成、カメラから直接、感情推定を行うシステムを作成していく。

5. おわりに

本研究では、基本6表情に現れるAUには含まれていない皺を含めた表情からの感情推定をするシステムを考案した。

ユーザの表情から感情を推定してコンピュータとのコミュニケーションに用いる場合、リアルタイムでの表情の認識が必要になる。本研究では皺の認識を追加して行っているため、認識の際に処理が多くなっているため、処理時間の短縮が必要である。また、AUの認識の際にニュートラル、最大、最小の値を一定値にして考えていたが、多くの人に応じて感情認識するためには、人それぞれにある各値を個人に対応したものを作出できるようにする必要がある。問題点を解決し表情からのより豊かな感情を推定できるシステムへと進めていく。

参考文献

- 1) Fujita, H. et al : Virtual Cognitive Model for Miyazawa Kenji Based on Speech and Facial Images Recognition, WSEAS Transactions on Circuits and Systems, Issue 10, Vol.5, pp. 1536-1543(2006)
- 2) Ekman, P. et al : The Facial Action Coding System; Consultation Psychologists Press, 1978
- 3) Ekman, P. et al : 表情分析入門; 誠信書房(1987)
- 4) 下田宏他 : 動的画像からのリアルタイム表情認識システムの試作; ヒューマンインターフェース学会論文, Vol.1, No.2, pp.25-32(1999)
- 5) Sebe, N. et al : Authentic Facial Expression Analysis; Sixth IEEE International Conference on Automatic Face & Gesture Recognition, pp.517-522, 2004
- 6) Tian, Y. et al : Recognizing Action Units for Facial Expression Analysis; IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.23, No.2, February 2001
- 7) Cootes, T. F. et al : Active Appearance Model; European Conference on Computer Vision 1998, Vol.2, pp.489-498, 1998

オノマトペに着目した賢治スタイルの抽出

Extraction the Kenji-style focused on onomatopoeias in Kenji works

インテリジェントソフトウェアシステム学講座 0312003090 高橋 一馬

指導教員：藤田ハミド 横松理樹

1. はじめに

現在、我々は「宮澤賢治プロジェクト」[1]と称し、コンピュータ上に宮澤賢治を再現することを目指している。本プロジェクトでは、実在の宮澤賢治のように、人と対話し、インラクティブな対応を実現するシステムの構築を目指している。そのような機能を実現するための要素の一つとして、賢治らしい話し方が挙げられる。ここで賢治らしいとは、賢治の思考や嗜好、思想などに従うことを意味しており、本稿ではこれを

「賢治スタイル」と呼ぶ。賢治スタイルを構築するためには、賢治本人に直接情報を得る必要があるが、賢治はすでに他界しており、不可能である。しかし、賢治が多くの作品を残している点に着目し、それから賢治スタイルの一端として、賢治らしい言葉・言い回しの抽出を試みる。賢治の残した作品は、童話や詩であり、話し言葉ではない。書き言葉と話し言葉の差はあるが、同じ人物が発した言葉という点では共通しており、その点から有益な情報を得ることが期待できる。

本研究では、賢治スタイルを抽出するために、賢治作品におけるオノマトペに着目した。オノマトペとは、「事物の音や人・動物の声、物事の状態や様子などが感覚的に音声化して表現する語」である。オノマトペは感覚的な言葉であることから、作者の感性、個性が現れることが考えられる。また賢治はオノマトペの使い方に対し評価が高く、文学の分野で研究が盛んに行われている[2]。このことから、賢治スタイルを抽出するためにオノマトペに着目することは有用であると考えられる。本研究では賢治作品に現れるオノマトペに対し、それが表す表現対象との組の出現分布を求め、それらを他の作品や他者の作品と比較し、特徴的な点を賢治スタイルとして抽出する。抽出された特徴を、コンピュータ上の賢治の発話内容に反映することで、賢治らしい発話の実現が期待できる。

2. オノマトペに着目した賢治スタイルの抽出

2.1 オノマトペ

オノマトペとは「擬音語・擬声語・擬態語を包括的にいう語」である。擬音語・擬声語とは「事物の音や人・動物の声などを表す語」であり、擬態語とは「物事の状態や様子などを感覚的に音声化して表現する語」である。このようなオノマトペを文章中で利用することで、臨場感を与えることや、様子や行動を強調することができる。また通常の言葉では表現しにくいことを感覚的にうつたえることができる。このため、オノマトペには作者の感性が反映されやすく、個性が現れることが考えられる。これまでの文学研究においても、オノマトペは研究対象であり[2][3]、その点からも重要な指標になるとされる。

2.2 推定手法の流れ

本研究で行う推定手法は、①オノマトペの抽出、②

オノマトペの表現対象の抽出、③出現傾向の抽出、④比較の4段階からなる。以下、それぞれの段階を説明する。

①オノマトペの抽出：

対象となる作品からオノマトペの抽出を行う。オノマトペについては、慣用的に利用されているもの他に、方言や作者独自のものもある。そのため、辞書などを参照し、オノマトペと考えられるものを手作業で抽出する。

②オノマトペの表現対象の抽出：

オノマトペは擬態・擬音語である。そのことから、各オノマトペが表現しているものが、文中もしくは文外に存在していると考えられる。それらに対し、文脈を意識しながら手作業で抽出し、オノマトペとの組を作成する。

③出現傾向の抽出：

②まで抽出したオノマトペと表現対象の組に対し、同じものをまとめていくことで出現傾向を取り出す。出現傾向としては、オノマトペまたは表現対象の一方を固定し、それに対するもう一方の出現回数に基づく出現分布を求める。このとき、文中に出現している語をそのまま用いるだけではなく、類似した特徴でグループ化を行い、そのグループ単位での出現分布も同様に求める。グループとしては今回は次のものを用いる。

(a) オノマトペに対するグループ：

(a) 音韻に基づくグループ化：天沼[4]は、オノマトペを構成する日本語音韻の並びから分類する手法を提案している。この分類を利用し、音韻により36種類のパターンにグループ化する。

(b) 音の種類によるグループ化：丹野[5]は、音素の種類により、清音、濁音、半濁音によってオノマトペを分類している。これを活用し、オノマトペを清音、濁音、半濁音、その他の4種類にグループ化する。

(c) オノマトペの直後の語によるグループ化：オノマトペにつながる語によってもオノマトペの使い方に特徴が現ると考え、これによりグループ化する。今回は、「句読点・と・の・する・して・その他」の6種類を考える。

(d) 表現対象によるグループ化：表現対象については、その内容について、日本語百科大事典[6]を参考に、①モノの状態・変化、②感情、③音、④声、⑤生物の動作、⑥様態、⑦感覚の7つにグループ化する。

オノマトペと表現対象に対し、文中での表現や以上のグループ化を行ったものに対し、それぞれ出現分布 $P(X)$ を次式で求める。ここで X はオノマトペまたは表現対象そのものかグループを示す。

$$P(Oi) = \frac{Num(Oi, Gj)}{\sum_{j=1}^N Num(Oi, Gj)} \quad (i=1 \dots M)$$

$$P(Gi) = \frac{Num(Oi, Gj)}{\sum_{i=1}^M Num(Oi, Gj)} \quad (j=1 \dots N)$$

ここで、 Oi はオノマトペまたはオノマトペの一つのグループ、 Gj は表現対象またはそのグループを示す。また M は Oi の種類数、 N は Gj の種類数を示す。 $Num(X, Y)$ は X と Y が同時に出現した個数を示す。

④ 比較：

複数の作家の複数の作品に対し、①から③の手続きで出現分布を求める。それらを比較したとき、ある作家の作品における出現分布が、その作家の作品中では類似しているが、他の作家の作品とは異なれば、それを特徴的な点として捉える。出現分布の比較方法としては次の方法をとる。①出現分布を確率分布と見立て、同じ確率分布に従っているかを検定する適合度の検定により、同一の分布か否かを判定する。②出現分布をベクトルとして捉え、それらのベクトルのなす角にもとづく類似度により評価する。一定の値以上であれば、類似していると判定する。

これらによって賢治作品を特徴づける出現分布を賢治スタイル、賢治の文章の特徴として捉える。

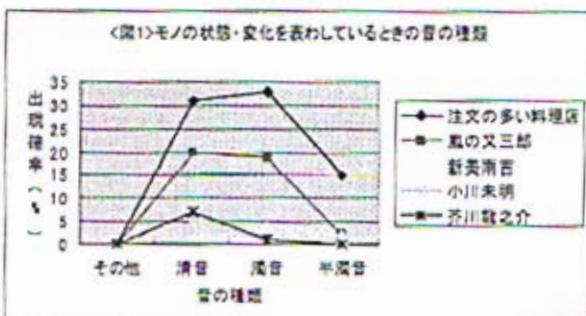
また得られた出現分布は、システムの発話内容にオノマトペを追加する際に利用される。これにより、賢治らしい発話の実現を目指す。

3. 評価実験

現在、本手法の有用性を評価するために、表1に示す、宮沢賢治の作品2点および賢治と同時代の作家である、新美南吉の作品4点、童話作家小川未明の作品4点、芥川龍之介の作品2点に対し、出現分布の抽出、比較評価を実施している。またオノマトペの抽出には、大阪国際児童文学館共同研究プロジェクトの成果[7]を参照している。

表1：実験に利用した作家と作品	
作家	作品名
宮沢賢治	注文の多い料理店、風の又三郎
新美南吉	手袋を買ひに、ごん狐、おじいさんのランプ、牛をつないだ椿の木
小川未明	野ばら、赤いろうそくと人魚 はてしなき世界、ある日の先生と子供
芥川龍之介	蜘蛛の糸、杜子春

出現分布の例として、作品・作家ごとの音の種類、表現対象の組を図1に示す。



現在、実験を進めている段階であり、実験の結果・考察については卒業論文に提示する予定である。

4.まとめ

本稿では、宮澤賢治の作品から、オノマトペに着目し、その利用方法から宮澤賢治の発話特徴（賢治スタイル）を抽出する手法を提案した。本手法は、オノマトペと表現対象との組合せに対し、他の作家と異なる点を特徴として捉えることを試みるものである。現在実験中であり、その結果を元に本手法の有用性を評価する必要がある。さらに、本提案手法の課題としては次のものが挙げられる。

- 開発労力の軽減：本提案手法では、オノマトペと表現対象の組の抽出を人手で行っている。これは、利用者に対し、多くの労力を要求することとなる。そのためコンピュータを活用することで、多くの作品を対象にできるという利点が発揮できない。形態素解析や構文解析を利用するといった機械化を行い、利用者の労力を軽減する必要がある。
- 比較観点の検討：現在の手法はオノマトペとその表現対象のみであり、範囲が限定されている。また表現対象については利用者（分析者）の判断に左右される点が大きい。コンピュータによる大量データ処理を考慮するのなら、利用者によって判断が変化しにくい観点からの比較を検討する必要がある。また今回はオノマトペを利用している場合に絞っているが、利用しない場合との比較も必要になると考えられる。

参考文献

- 岩手県立大学「宮沢賢治プロジェクト」：
<http://www.fujita.soft.iwate-pu.ac.jp/KENJI/index.html>
- 田守：「宮澤賢治のオノマトペ」、影山、岸本編・日本語の分析と言語類型、くろしお出版(2004)
- 滝浦：「オノマトペ論—ことばにとっての〈自然〉をめぐる考察ー」、共立女子短期大学文科紀要、第36号(1993)
- 天沼：「擬音語・擬態語辞典」、東京堂出版(1974)
- 丹野：「オノマトペ《擬音語・擬態語》を考える」、あいり出版(2005)
- 金田一晴彦他：日本語百科大事典、(大修館書店)(1995)
- 上田信道他：宮沢賢治『注文の多い料理店』データ集：
http://www.d1.dion.ne.jp/~ueda_nob/authors/pcken/pcken.html

韻律に基づく音声からの感情推定に関する研究

Estimation of Emotion in Speech Based on Prosodic

インテリジェントソフトウェアシステム学講座 0312005304 澤井 夏美

指導教員：藤田 ハミド 柏松 理樹 羽倉 淳

1. はじめに

現在、我々は「宮沢賢治プロジェクト」¹⁾と称し、コンピュータ上で宮沢賢治を再現するプロジェクトを実施している。本プロジェクトでは、人と同じように対話することができる仮想空間上の宮沢賢治の構築を目指している。人と同様に対話するには、人と同様に対話時に相手の表情や仕草、音声から感情を推定し、それを考慮する必要がある。本研究ではこのような機能の一部として音声からの感情推定に着目する。

音声は大きく、個々の音素である音韻と、音素の強弱・高低・長短などの韻律とに分類される。主に感情が表出されるのは、韻律である²⁾。そのため、本研究では、韻律からの感情推定を試みる。このような韻律から感情を推定する研究は、現在までにいくつか実施されてきている³⁾⁴⁾。しかし、明確な推定手法が確立されていないのが現状である。

以上のような背景から、本研究では、既存の研究をベースとし、独自の観点を加えた韻律からの感情推定を行う新しい手法を提案する。本手法は、音の強弱・高低・長短といった音声が持つ韻律情報である韻律特徴とそれに対応付けられた感情からなる感情推定ルールを用い、音声から感情を推定する。感情推定ルールにおける韻律特徴としては、従来の感情推定手法で用いられている強弱（パワー）・高低（ピッチ）・長短（発話時間）の代表値に、その変化を加える。また処理対象ではなかった間（ま）に着目する。さらにこれらの特徴の値を、自然な発話から得られた韻律特徴によって修正する。これらにより、従来の手法よりも自然な発話における推定精度の高い手法を実現することが期待できる。

2. 韵律に基づく感情推定手法

図1に示すように、本手法は、発話データから構築された感情推定ルールを用いて、音声から感情を推定する。感情推定ルールは、音声の韻律特徴とそれに対応づけられた感情から構成される。以下、発話データおよび感情推定ルールの構築方法について述べる。

2. 1 発話データ

発話データとしては、例えば演劇の台詞のような「ある感情を伝えることを目的とした発話」である意図的発話と、例えは会話のなかで発せられるような自然発話の2種類の発話データを用いる。本手法では、発話における間（ま）も特徴として捉

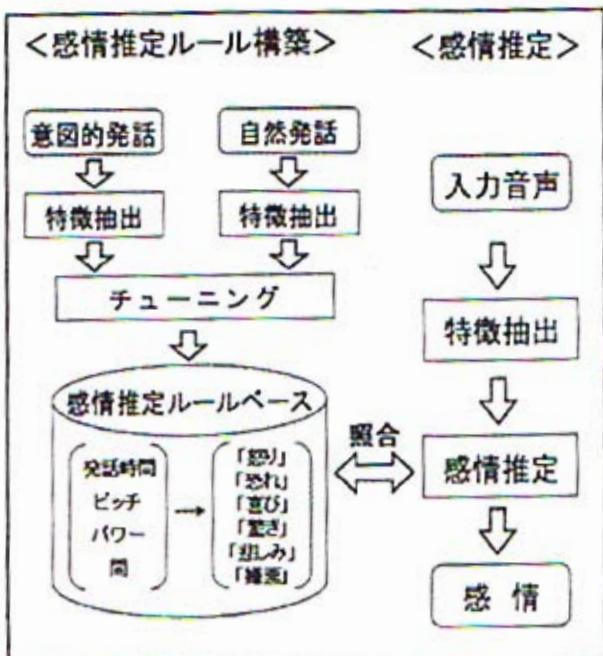


図1：韻律に基づく感情推定手法概念図

えるため、発話は一語句単位ではなく、一定以上の時間長を持ったデータを利用する。

それぞれの発話に対しては、Ekmanの研究⁵⁾をベースに、「怒り・恐れ・喜び・驚き・悲しみ・嫌悪」の6感情のラベルを発話者自身が付与する。このとき、一つの発話に対し、複数の感情ラベルをつけることも許す。

意図的発話の収集は、発話者に対し、台詞を提示し、前述の感情のいずれかを意識して発話してもらうことで収集する。これは基本的に従来の手法と同じである。一方、自然発話に対しては、より自然な状況かつ多くの感情が表出する発話をとる必要がある。従来の研究では、ある話題に対する対話を用いることが多かったが、感情が見えにくく、また表出される感情も少ない。そのため十分な自然発話の収集が難しい。そこで、本研究では、他者と交渉を行う必要があるゲームを発話者数名で実施してもらい、そのときの発話を収集することを行う。交渉ゲームにおいては相手との駆け引き、成功不成功など様々な状況が生じるため、感情が表出しやすく、また多様な感情の表出が期待できる。

2. 2 感情推定ルールの構築

(1) 意図的発話からの韻律特徴の抽出

はじめに意図的発話に対し、韻律特徴を取り出す。

韻律特徴としては、周波数分析によって得られるものと、間(ま)に対するものがある。周波数分析によって得られる特徴としては、パワー、ピッチ、発話時間を利用する。これらの値に対し、一発話における最大値、最小値、平均値を捉える。これらは従来の手法でも用いられているものである。本手法では、これらに加え、音素単位の変化量、変化のパターンについても特徴として捉える。従来の手法が、言わば点として捉えていたのに対し、本手法は、線でもとらえることを試みる。これは、人の感情は韻律の変化に現れるという知見²⁾に基づくものである。

また、従来の方法では処理のタイミングであつた間に対しても処理を加える。一発話内に含まれる間の回数、平均長、割合についても特徴として捉える。間を特徴とみなす理由は、人が話す場合、感情によって間も左右されるという知見²⁾に基づいている。

(2) 自然発話からの韻律特徴でのチューニング

次に意図的発話と同じ感情でラベル付けされた自然発話に対し、意図的発話の時と同様の方法で韻律特徴を抽出する。この結果と(1)の結果とを比較し、合成することで、チューニングを行なう。現時点では合成としては、それぞれの値の重心をとることを行う。

(3) 韵律特徴と感情の対応づけ

(2)で取り出した韻律特徴について、変化の傾向が共通するものをまとめる。まとめる方法として回帰木の手法を活用する。値ではなく変化の傾向とする理由は、個人により基本となる値が異なるためである。まとめた韻律特徴に対し、そこに付与された感情を対応づける。この時、同じ韻律特徴を持つものに対し、複数の感情が与えられていることを考慮し、それぞれの割合を感情ごとの強度とする。

2.3 推定手法

構築した感情推定ルールを用い、感情の推定を行う。はじめに新たに入力された発話に対し、感情推定ルール構築時と同様に韻律特徴を抽出する。次に韻律特徴と推定ルールとを照合し、類似しているものを取り出す。最後に取り出したルールの感情部を合成し、推定する感情として取り出す。感情の合成については、各感情の強度の平均値をとることで実現する。

3. 実験

本手法の有用性を評価するために実験を実施している。実験の内容は、発話データの収集、感情推定ルールの構築および推定実験である。

3.1 発話データの収集

はじめに意図的発話データとして、20歳前後の男子4名が、20個の台詞を、6感情それぞれを込めて発話したものを収集する。このとき音声のサンプリング周波数は16kHzとした。

次に自然発話データとして、意図的発話の発話者が、交渉を伴うゲームを実施している際の発話を記録する。また、収集した発話に対し、発話者自らが6感情のラベルをつける。図2に自然発話

において出現した感情の種類を示す。感情ごとのばらつきはある多くの感情が表出していることが見て取れる。

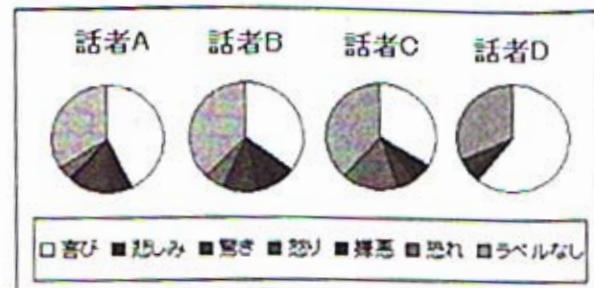


図2：自然発話における感情の種類

3.2 ルール構築と推定

集められた発話データに対し、韻律特徴を抽出し、ルールを作成する。なお自然発話における一発話の単位としては、固定時間(30秒)を用いるほか、同一のラベルがつけられた発話、間から間の間などいくつかのパターンで試す。これらに対し2.2で示した手法で構築したルールを用いて、推定を行なう。推定の評価は、処理結果と発話者の回答とを比較し、その一致度で判定する。また従来手法との比較を行う。

4. おわりに

本研究では、音声の韻律特徴から発話者の感情を推定するため、音声の強弱・高低・長短の代表値に加え、それらの変化、間に注目することにより感情推定を試みる手法を提案した。

また、本研究で得られる韻律特徴に関する成果は、宮沢賢治プロジェクトの機能の一部である音声合成部においても有用であると考えられる。音声合成の際、強弱、長短、高低などの韻律特徴を本研究で構築する感情推定ルールに基づき変化させることによって、より自然な感情音声の合成が可能になると期待できる。

今後の課題としては、実験を通じ、本手法の有用性を評価するとともに、韻律特徴の検討、推定精度・処理速度の向上が挙げられる。

参考文献

- 1) 宮沢賢治プロジェクト：
<http://www.fujita.soft.iwate-pu.ac.jp/KENJI/index.html>
- 2) 小池生夫他：“応用言語学辞典”，研究社，(2003)。
- 3) Pierre-Yves Oudeyer：“The production and recognition of emotions in speech: features and algorithms”，International Journal of Human Computer Interaction, 59(1-2):157-183, (2003)。
- 4) 飯田朱美、伊賀聰一郎、安村通晃：“音声にみられる感表現について-聴取実験結果の分析”，音声言語情報処理学会, (1997)。
- 5) 前川喜久雄、北川智利：“音声はバラ言語情報をいかに伝えるか”，認知科学, Vol9, No.1(2002)。
- 6) P.Ekman, W.V.Friesen, 工藤力 訳：“表情分析 入門-表情に隠された意味をさぐる”，誠信書房, (1987)。