

機械の心（感情認識から心理解析まで）

光吉 俊二^{††} 任 福継[‡]

[†]株式会社エイ・ジー・アイ

[‡]徳島大学

心理学において、人間の「心」では、深層にある「情動の要因」から、表層としての「感情」が現れるとされている。それを、自我や超自我と呼ばれている抑制機能により、コントロールして、人間らしい社会関係や心理思考を行っていると考えられている。これは、脳科学においても実証されだしてきた。ここで、機械に「心」と「情報の因果関係」の関係をルールとして与えると、簡単な心的装置として一種の価値判断を構築することが可能になると考える。ここでは、「感情認識」から「心理分析」の一部を再現することができるのかを考察する。A.規格化（固定）されたルールにおいて認識する手段。B.固定された基準からの「ブレ」や「ぼらつき」から、人間の「揺らぎ」の特性を知り、規格化されたルールの個人補正（個人の心理特性や属性に基づく）を行い、より高度な認識を行う手法。以上2つの概念を用いて、機械の「心」として提案する。

Mindo of Machine

Shunji Mitsuyoshi^{††} and Fuji Ren[‡]

[†] AGI Inc.,

[‡] The University Tokushima

In psychology, human's mind is assumed to be appearance of "emotion as the surface" from "Factor of deep emotionality". Human is assumed to control emotions by "Control function that is called the ego and a superego", and to do a human idea. Moreover, it is smoothly related to the society. This fact has been proven in the brain science. Then, "Causal relation between the mind and information" is given to the machine as a rule. It is thought that the machine becomes possible the construction of an easy value judgment as the psychic apparatus. Here, whether a part of "Psychological analysis" can be reproduced from "Emotion recognition" is considered. A.Means to recognize in rule standardized (fixation). B.The characteristic of human's "Swing" is recognized from "Blur and difference" from the fixed standard. And, more advanced recognition is done by correcting the individual of the standardized rule. It proposes it by using two concepts above as "Mind" of the machine.

1. まえがき

ロボットや機械が家庭や社会と共存するためには、スムーズな機械とのコミュニケーションが必要となる。しかし、この分野の研究は発展途上であり、音声認識や言語理解などの研究と比較しても、明快な解が出ているわけではない。そこで、ロボットや機械の「心」を再現しようとした場合、2の課題の克服が必要になる。その一つが、機械の感受性である。この実現には、優れた音声認識や感情認識が必要になる。もう一つは、感情情報からの心理特性の察知である。これには、知能や心理思考と呼ばれる機能が必要になる。しかし、音声認識や知識判断の研究は進んでいるが、情動や心理思考の方面の研究はこれからである。特に、人間の感性のズレのような、抽象的なテーマになる場合、心理学や臨床精神分析学で使われる仮定や仮説を利用する必要がある。しかし、これらを工学的なモデルとして扱う上で、科学性に充分注意しないとイケない。そこで、抽象的で「不確実」に「揺らぐ」対象を、「相対的」に捉えて、一つの固定された手法で規格化し、そこから来る情報の「ぶれ」を検知する確実な手段

が必要になる。

1. 音声の波形特性からの感情認識
2. ロボットコミュニケーションに必要な心理系技術による心理分析（心理学仮説の利用）

以上の要素技術を再現性、信頼性に充分注意しながら、ロボットや機械が持つ「心」の一つの例として提案し、最新の脳科学の研究と比較した。

2. 音声からの感情認識

筆者は最初に「音声からの感情認識」を開発することにした。何故音声から始めたかであるが、視覚情報のパタンマッチングは難しいからである。音声の場合、その聴覚器官の研究が昔から行われていて、大体、数学モデルで記述でき、再現できる。また、周波数や音量（パワー・エネルギー）といった少ない要素で判断しているという便利さがある。音声からの感情認識研究は、音声からの言語情報認識（音声認識）と比べて未発達と言ってよい。そのため、基準となる人間自身の音声からの感情認識の仕方や、認識性能を知る必要があった。また、感情要素抽出のため、音声波形の解析

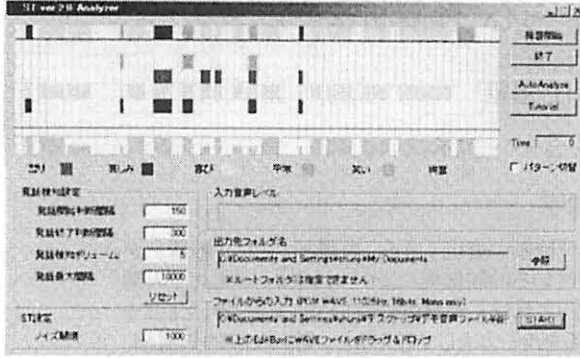


図1 多感情同時認識のビューアー

から始める必要があった。音声を構成する要素は「パワー」「周波数」と、それらから構成された「フォルマント」や時間遷移により変化する抑揚情報である。音声研究の黎明期から、基本周波数 F0 と感情の関係は指摘されていた。しかし、その「感情」の基本周波数 F0 を明確にピッチ情報とし、感情認識をさせた手段を開示している資料がなかった。

- 音声認識を行わないで、音声波形からの基本周波数 F0 の抽出と、パワーの遷移情報で得られる特徴を条件式群とパラメータで判定する手段による、「音声の抑揚からの感情認識」を作った（詳しくは、文献1を参照）。（パラメータ導出の要素は表1参照）
- 複雑な人間の複合感情にも対応するべく、同時に複数の感情の出現を検知させる手段と、その中から、第一候補の感情を選別する手段を作った。（図1参照）

2.1 感情認識の特徴

感情認識では、細かい発話区間での F0 やパワーの変化を計算し、（詳しくは、文献2-4を参照）全体の発話区間としての変化をさらに計算する。この計算手法により、「全体的にこんな雰囲気」という計算が可能になる。このようにして得られたパラメータ値に対して、データマイニングによる決定木を通すことで感情判定を行う。データマイニングでは、人間自身による気の遠くなる作業と、マイニングツール（詳しくは、文献5を参照）での作業を比較しながら、交互に行っている。

決定木はあらかじめ作成した IF 文の集合である。よって、IF 文なので判断基準が明確であり、「なぜそのような判定に至ったのか」が、解析できる判定ロジックができるようになった。これは、心理解析に最も重要な要素となる。簡単に説明すると、人間の「揺らぎ」を確実に計測しようとする、固定された基準を設けて、そこからどれだけ「揺らぐ」のかをみる必要がある。そして、この固定された基準に判定ロジックを用いるわけである。

2.2 感情認識と人間の認識比較

表3,4,5（文献1から引用）は人間同士の感情認識と感情認識との比較である。試験：人間の音声からの感情認識能力を知るために、本人の主観を基準とし、感情認識との比較

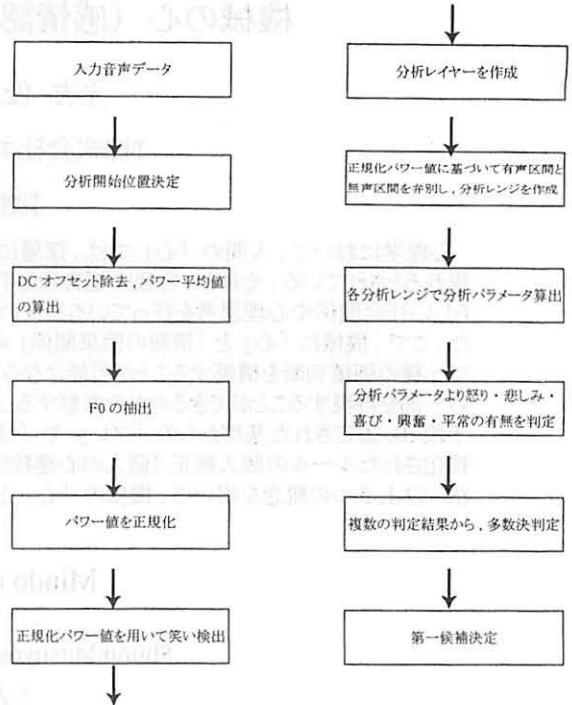


図2 感情認識の流れ図

表1 感情認識判断要素

要素	状態
正規化パワー値	パワー値をパワー平均値で割った値
F0 の分析	基本周波数の分析

表2 音声規格

11.025KHz16 ビット・モノラル、線形 PCM wave
ファイル数
分離音声資料総数 29876（発話者総数：約 200 名）

を正答率として比較した。

- 自然に発話された本人の主観との正答率との比較をした。
- 自然に発話された資料での発話した本人とは別の第三者6名の主観の正答率との比較をした。
- 日本語の解らない外国人（中国3名、米国3名）による日本語の感情発話の外国人の主観評価の正答率を言語に依存しない場合とした。
- 主観評価を1週間後、同じ本人が再評価して、一致しなかった率を人間の主観の揺らぎとした。

2.3 感情認識の問題点・人間の不確実性

感情認識は、人間の主観がそのまま判定手段に影響し、固定されているため、「感情の鼓膜」のような機能になる。人間自身の主観のズレがそのまま主観評価と認識結果の一致に影響するため、感情認識率 100% の実現障害となっている。本人の主観評価実験では、時間を置くと2割から4割も一致しない（表5参照）という事実もあり、感情認識率 100%

表3 言語理解がある場合の自然発話での人間主観との一致率

試験の特性	感情認識
l. 本人主観 (日本人)	69.75% (音声認識なし)
m. 第三者主観 (日本人)	60.71% (音声認識なし)
平均感情認識率 (日本人)	62.19% (音声認識なし)

表4 言語依存しない場合の人間同士の主観一致

試験の特性	人間
n. 第三者主観 (日本語を知らない外国人の場合)	(日本語の感情発話の正答率平均) 55%

表5 人間自身の主観の揺らぎ

試験の特性	人間
o. 時間経過による主観の揺らぎ	20.01 ~ 26.01%

表6 感情別認識率特性

	自然感情発話 (Files)	意図的感情発話 (Files)
怒り	52.54% (59)	70.29% (340)
哀しみ	35.00% (20)	52.17% (368)
喜び	65.38% (52)	53.33% (390)
平常	92.86% (56)	69.86% (511)

を実現することは難しい。感情認識単体で認識精度を上げてすぐに限界を迎える。また、誰が聞いても「感情が伝わらない」発話の場合は、感情認識でも認識が低い。

3. 心理学仮説の利用による不確実性回避の構想

人間は、初対面の場合、瞬時にその人がどういう人なのか？を知ることは難しい。特に、電話などの場合、顔が見られない状況があり、即時に相手の性格や性向を知る手段がない。「音声からのリズム感での感情認識」でもこれは同じである。しかし、身内や親しい間柄では「阿吽」の呼吸のように「相手の気持ちを性格、性向に応じて判断」する。これは、人間同士の付き合いが永く、相手の発話特性をフィルタリングできているからである。

そこで、筆者は感情認識において得られた感情要素（情動要素）を基にして、時間軸での変化量（時制における遷移情報）などから「相手の気持ちを性格、性向に応じて判断」して、感情認識の判定ロジックやパラメータ閾値をキャリブレーション（補正）し認識を上げることが可能であると考えられる。

これを心理系技術 (PST) とした。(図3参照)しかし、PSTの実現には、感情判定のロジックが、発話者の心理状態に対応して、物理考察できる状態でないといけない。

3.1 PS ⇔ D 振動 (Move) 原理の心理特性

PS ⇔ D 振動 (Move) 原理とは、心理思考のことである。(詳

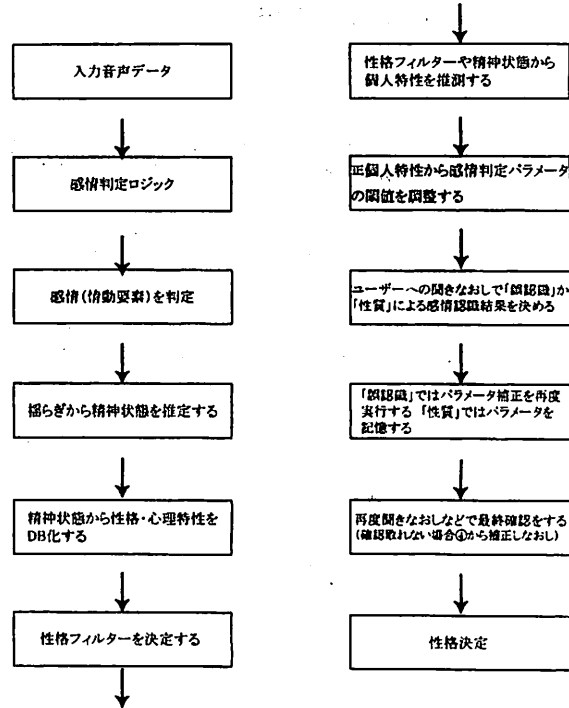


図3 PSTによる個人補正手段

しくは、文献7を参照)これは、自我判断とも連動するため、心理思考と自我の境界を明確に分けることは難しい。しかし、人間の精神的活動から思考の原理を説明する具体的な手段である。これは、PSTの最も重要な機能となる。PS ⇔ Dは、「人間は、大きな不安に直面した時に妄想分裂と抑鬱の間を揺らぎ（振動）ながら耐える」というメラニー・クライン理論である。(詳しくは、文献6を参照)要約すると、人間には、問題、トラブル、アクシデントに直面した場合、妄想（大変なことになった）分裂（どうしよう・動揺）< PS=Paranoid schizophrenia >と抑鬱（ひたすら謝ろう・現実逃避）< D=Depression >の間を揺らぎながら耐えるという、心理特性を指す。これを原点として、Wilfred Bionが進歩させ、思考や心の判断の基本メカニズムとして、妄想分裂と抑鬱の揺らぎにより、人間は思考統合すると説明した論理である(詳しくは、文献7を参照)。PS ⇔ D 振動 (Move) 原理はコンピュータの感性ユーザーインターフェース < SUI > (図4参照・日本 SGI 社共同開発)や通信、自動車、家電、コールセンター、ゲームなどへ、すでに導入が進められようとしている(詳しくは、文献8を参照)。

3.2 脳科学と PS ⇔ D 振動の類似性

「感情」は人間の心の表層に現れた「現象 (シンボル)」ともいえる。しかし、そのシンボル化された「感情」には、何から起因するのか、原理、要素、理由といった原因が存在する。それを「深層」すなわちシンボル化以前の状態として扱う。そのため、脳の構造やシステムの理解や再現が重要になってくる。最終的に対象となる、人間の脳自体が完全に演繹説明 (数理記述) されていない状態では、脳科学との密

心理分析の構造

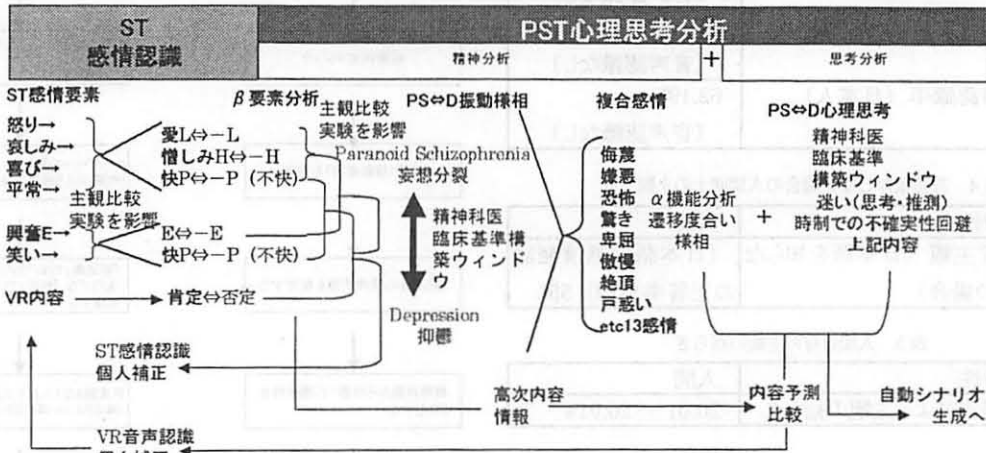


図4 PS ↔ D 振動 (Move) 原理と PST モデル (SSC)

接なコミュニケーションで研究が進められることが必要になる。ただし、表層に現れた「感情」の因果関係に限れば、シンボル化されているためプログラムやアルゴリズムとして記述可能である。また、感情と「深層」とは密接な因果関係があると考えられる。感情と「深層」の因果関係の物理考察は、実際の脳の中でのでき事との比較となる。

3.3 無意識からくる深層と感情の関係

NICT(独立行政法人・情報通信研究機構)の調査では、<視覚的意識が自発的に変化するときには、脳が離散的・確率的に振る舞ったり、化学反応と同じ速度論に従ったりすることをきちんと定量的に示すことができる>ことを発見したようである。「化学反応と同じ」というのは、脳がデジタル回路のような整然としたロジックに従って動いているのではないことを示唆する。つまり、熱運動している分子同士がランダムに衝突して化学反応を起こすように、脳も、自分勝手に振る舞う膨大な数の神経細胞を基本としているが、それが脳全体として機能するときには<何らかのうまい「統合」の仕組み(シンクロニゼーションなど)>があって、一見整然と振る舞っているように見える、ということである。

ところが、基礎にある神経集団がランダムな過程は、非常に本質的な場面で整合性のある答を出すために、機能している。それが柔軟性とか創造性といった、従来のコンピュータでは再現が難しい、脳ならではの機能の基盤を与えているのではないかと考えている。統合後の状態は意識化されるので本人にもわかるが、そこに至るプロセスは意識にのぼることがない。(ランダム過程をそのまま意識化すると行動上の混乱

が生じ、その生き物が自然界で生き残れなくなるからだと推定している。)そのため、「ひらめいた」結果はわかるのだけれど、なぜそれをひらめいたのかは本人にもわからない、というようなことがあるのだと推定する。以上が、研究者の見解である。

4. まとめ

「深層」統合以前の状態と「意識」(シンボル)状態の関係が、「ひらめき」や「アイデアの源泉」の解明につながると思われる。それは、

- (1). 熱運動している分子同士がランダムに衝突して化学反応を起こすように、脳も、自分勝手に振る舞う膨大な数の神経細胞を基本としている。
- (2). (1)により、脳がデジタル回路のような整然としたロジックに従って動いているのではない。
- (3). (1)が脳全体として機能するときには<何らかのうまい「統合」の仕組み(シンクロニゼーションなど)>があって、一見整然と振る舞っているように見える。
- (4). 基礎にある神経集団のランダムな過程は、本質的な場面で整合性のある答を出すために、機能している。
- (5). 人間独特の柔軟性とか創造性の由来をここに見ることができるかもしれない。

以上が NICT において、現在研究されている内容をまとめたものであるが、(1)と(2)は脳科学や医療により確認されている事実として、受け止めていい。そして、(3),(4)こそが、重要な示唆を筆者にあたえた。(3)の<何らかのうまい「統合」の

仕組み>を脳情報の研究者は(シンクロニゼーションなど)と説明している。これは、精神分析学でいるPS⇔D振動や「心」の共鳴といった、シンボル化された「もの」の「揺らぎ」「統合」現象に似ている。(4)の<ランダムな過程>は筆者が2003年のIEEEで発表した心量子の確率モデルの振舞いに似ている(詳しくは、文献9を参照)。また、(4)の<本質的な場面で整合性>は超自我などでの抑制やD凝縮にも似ている。

謝辞

この分野において、筆者のようなものでも発表の機会が巡ってくる時代になったことに感謝いたします。同NICTの長尾理事長、および脳情報の研究者のみな様にもご理解を頂き、ご支援ありがとうございます。そして、NTT持株会社の井上第三部門長、NTTコミュニケーション科学研究所の片桐所長には感情認識へのご理解とご指導を頂いており、心から感謝いたしております。筆者の仮説にご賛同頂き、スタンフォード大学での研究にご援助いただきけた、MIT名誉教授エドガー・H・シャイン先生にも感謝の意を送りたいと存じます。

参考文献

- [1] 光吉俊二, 任福継, [人間の感情を測定する] 電気学会誌 解説論文 掲載予定 (2005).
- [2] 国枝伸行・島村徹也・鈴木誠史:「対数スペクトルの自己相関関数を利用したピッチ抽出法」, 信学論, vol.J80A, no.3, pp.435-443, (1997).
- [3] 吉尾重治・趙 奇方・島村徹也・鈴木誠史:「平方根及び4乗根パワースペクトルの自己相関に基づくピッチ抽出」, 信学論, Vol.J84-A, no.3, pp.436-440, (2001).
- [4] Jouji Suzuki, Miyuki Setho and Tetsuya Shimamura: "Extraction of Precise Fundamental Frequency based on Harmonic Structure of Speech," Proc. 15th Int. congress on Acoustics, pp.161-164, (1995).
- [5] Data Mining Tools C5.0 RULEQUEST RESEARCH (1997).
- [6] "The Writing of Melanie Klein (Vol.3). Envy, and Gratitude and Other Works(1946-1963)". 日本語訳, 小此木啓吾, 岩崎徹也: "メラニー・クライン著作集 1946-1955 4妄想的・分裂的世界" 誠信書房 (1985) .
- [7] Joan and Neville Symington "The Clinical Thinking of Wilfred Bion" Routledge, London, (1996).
- [8] 井上友二:「レゾナントコミュニケーション環境の創造に向けた研究開発」, NTT 技術ジャーナル, Vol.17, No.4, pp.14-25 (2005).
- [9] Shunji Mitsuyoshi and Fuji Ren: "The Sensibility Inference Function by Psycho-Quantum Computer," IEEE International Conference on System, Man & Cybernetics, pp.1679-1686, Washington, D.C., (2003).