

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年5月18日現在

機関番号：16101

研究種目：挑戦的萌芽研究

研究期間：2009～2011

課題番号：21650030

研究課題名（和文） 工学的側面から心の豊かさを扱う学問体系の確立

研究課題名（英文） Establishment of an academic system of Enriching Mental Engineering

研究代表者

任 福継 (REN FUJI)

徳島大学・大学院ソシオテクノサイエンス研究部・教授

研究者番号：20264947

研究成果の概要（和文）：豊心工学の定義、学問体系の研究を行い、工学としての心の豊かさを定義した。即ち、心の豊かさは人の命存在の価値観、そして外部からの様々な刺激に対する受ける能力を測る尺度である。ブログを収集し、タグ付けを行い、感情コーパスを構築した。この感情コーパスをもとに、感情の変化を時系列的に解析し、うつ病との関連性はある程度見つけて、自殺傾向をブログデータから推定できるアルゴリズムを開発した。

研究成果の概要（英文）：An engineering discipline that deals with richness of people's minds, in other words, Enriching Mental Engineering (EME) was proposed. In particular, its research content, methodologies, and tasks were described with respect to the academic framework of EME. Based on the results of psychological experiments, a mental state transition network was constructed using statistical methods. We proposed a method for estimating people's mental states and studied the emotional energy function. Further, we constructed a blog emotion corpus and proposed a method of detecting depression from blogs.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	1,500,000	0	1,500,000
2010年度	1,000,000	0	1,000,000
2011年度	700,000	210,000	910,000
年度			
年度			
総計	3,200,000	210,000	3,410,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：知的エージェント、感性情報学、心豊かさ、心的状態遷移、感情認識、感情状態、言語解析、豊心工学

1. 研究開始当初の背景

(1) 自殺者とうつ病

日本の自殺者の総数は、1998年以来連続で3万人を超えている。自殺予防総合対策センターのウェブサイトによると、

中高年と高齢者の自殺死亡率が減少している一方で、若年者のそれは増加傾向にある。これは、金銭や物質の不足よりも心の貧しさが自殺の主な原因ではないかと思われる。

(2) 工学はどのように貢献できるか

インターネットの普及に伴い、多くの人々が自分の心の内をブログという形で表現し、公開している。それらの中には、自殺に至る心の動きが記述されたものも存在する。例えば、元TBSアナウンサーが2008年5月26日に自殺した。彼女が死の直前まで綴っていたブログには「仕事の合間が一番苦痛です」「言葉が・・・出ないので」と精神的に不安定だったと思わせる記述が存在した。また、四川大地震で息子を失った原幹部は、災害復旧の責任者として、頑張ったが、後ほど自殺した。自殺前までのブログには「愛する息子へ・・・天国で会おう」、「さようなら、愛する妻・・・」などのメッセージが残されている。このような記述の内容を工学的な手段を用いて解析することで、記述者自身の心の状態を推測し、自殺傾向の有無やその他精神的な病気の有無などについて判断することが可能であると考えられる。更に、ブログの記述者に対してどのような外的刺激(アドバイスなど)を与えればよいか、といった「癒し」の方策の提示も可能であると考えられる。また、音声や顔表情、行動からも工学的手段である程度、人の心的状態を推定できると思う。

2. 研究の目的

豊心工学の学問上の定義、内容、手法を明らかにする。対象者が記したテキスト情報、表情や行動、声色などの情報から心の豊かさを定量的に測定する、ということを中心的な技術と考え、工学的方法を用いた測定法を確立する。特にうつ病や自殺傾向の検出に焦点をあて、心の状態変化を計測する方法を開発する。また同時に、測定された心の状態をより豊かに改善するためにどのような外的刺激が有効か、といった技術についても体系化する。

3. 研究の方法

本研究の目的は、ブログから自殺傾向の検出、及びその状態を回復させるための工学的な技術の開発を通して、人の心の豊かさを扱う工学、即ち、豊心工学の学問体系を構築することにある。この目的を達成するために、次のような研究方法で研究を行った。

- ① 心理学的な実験結果をもとに統計的な手段を用いて、「心的状態遷移ネットワーク」を構築し、人の心の状態を推定する方法を提案すること
- ② 大規模なブログを収集し、ブログに記述された内容や表現と、自殺傾向やうつ病との相関関係を解明し、テキスト情報から心の状態を推定する方法を確立する

こと

- ③ 自殺傾向が強い状態の人などを回復へと向かわせるために、適切な外的刺激(アドバイス等)を自動で選択するための方法論を確立すること
 - ④ 声色などの音声情報や、表情などの画像情報、また脳波や脈拍といった生理現象と心の状態の間の関連性と相互作用のメカニズムを推論し、様々なモードの入力をあわせて心の状態を推定する方法を確立すること
- 豊心工学の理論的な定義、学問とする方法論、内容と体系を確立すること

4. 研究成果

① 豊心工学学問体系の確立

(1) 豊心工学

豊心工学とは、様々な現代技術をもとに工学的側面から心の豊かさを扱う学問体系である。「豊心工学」では、対象者の生理的な情報、さらに対象者が記したテキスト情報、表情や行動、声色などの情報をもとに、心の豊かさを定量的に測定する、ということを中心的な技術としている。また同時に、測定された心の状態をより豊かに改善するため、どのような外的刺激が必要とされるか、といったことも体系化する。

豊心工学は、計算機科学、人工知能、計算言語学、臨床心理学、精神医学、宗教などの分野を横断し、融合させた新しい学術分野であると考えられる。

(2) 心の豊かさの工学上の定義

心を豊かにさせることに関しては、古くは宗教、最近では臨床心理学等の分野でいくつかの理論が提案されている。心の問題は従来、意識や身体機能の問題として考えられているが、本研究で提案しようとする「豊心工学」では異なる捉え方をしている。

定義1：心の豊かさは人の命存在の価値観、そして外部からの様々な刺激に対する受ける能力を測る尺度である。

工学上、心の豊かさは次の式(1)で表示される。即ち、心の豊かさ m はハード的健康要素 h 、ソフト的健康要素 em 、所在する個人的社会的環境要素 ev 、所处する人間関係要素 r の関数である。

$$m(p) = f(h, em, ev, r) \quad (1)$$

心の豊かさの外在表現は次の感情状態 e で描写できると考える。

$$e(p) = g(L, S, F, A) \quad (2)$$

ここに、 L は言語表現における感情であり、 S は音声における感情であり、 F は顔表情における感情であり、 A は行動に

における感情である。即ち、感情状態は人の言葉遣い、音声、顔表情及び行為の関数である。

定義2：心の豊かさの尺度はある期間における感情状態の時系列関数である。ここに、心の豊かさの尺度は1と-1の間の数字で表現される。即ち、心の豊かさの尺度 v は感情状態 e と時間周期 T の関数である。

$$v(p) = k(e, T) \quad (3)$$

$v(p)$ は心の豊かさの尺度を示す指標であり、高いほど心が豊かであるが、マイナスになると、心が貧しくなる。

(3) 豊心工学の課題

豊心工学の学問体系の確立には多くの課題があげられるが、大きく分ければ、次の2つがあげられると考える。

課題1：心の豊かさの尺度を計測する技術の研究

課題2：心の豊かさの尺度を向上する工学的手法の開発

② 心的状態と心的状態遷移ネットワークの構築

人の心の感情は幾つかの状態に位置し、情報処理過程に対応するいくつかの離散の状態の間を遷移すると仮定するが、ここでは、これを「心的状態」と呼ぶ。特に区別のない場合、心的状態と感情状態と同じ意味で使う。

人間の心的状態は、ある条件で、一つの状態から別の状態へ遷移することができる^[2]。状態間の遷移は同じ確率ではないが、外界の要因を無視すると一定の期待値が存在する。大量のデータを分析し、人間の個性情報に基づいて、心状態遷移ネットワークモジュールを構築することができる。

我々はすでに心理的なアンケート手法で、心的状態遷移ネットワークを構築するため、次の実験を行った。まず、予め7種の感情状態の遷移表 $S(7, 7)$ を作成した。外部刺激がない条件で、状態 i から状態 j に遷移する可能性を被験者により1から10までの数字で記入してもらった。数字の大きさは遷移の可能性の大きさと一致する。そして、外部刺激がある場合に、同じ実験を行った。

③ 感情エネルギー関数による外部刺激の測定

心的状態遷移ネットワークについて、外部刺激の種類により、その遷移確率は当然異なる。通常、外部刺激とは、心以外の外界からの広い意味上の情報と考えているが、本稿では、外部刺激は心的状態遷移ネットワークに作用するエネ

ルギーとする。例えば、嬉しい情報を受けて、その人の「話し」や「身振り」や「表情」などから「喜び」という外観感情を表示したものを「喜び」という外部刺激とする。本研究では、外部刺激を感情エネルギーと呼ぶ。即ち、人間がある感情状態にある場合に、感情エネルギーがある一定の閾値を越えた場合に、他の感情状態に遷移する。

定義3：感情エネルギー関数は人の言葉遣い、音声、顔表情、生理的な情報、そして行為等の要素によってその時点で人の感情刺激を計算する関数である。

ここに、感情刺激は心的状態遷移ネットワークに作用する外部刺激である。感情エネルギー関数は式(4)で示される。

$$E = E_e(L_e, S_e, F_e, P_e, A_e) \quad (4)$$

E は外部刺激、即ち感情エネルギーで、 E_e は感情エネルギー関数、 L_e, S_e, F_e, P_e, A_e はそれぞれ言葉遣いに基づいて推定された感情、音声に基づいて推定された感情、顔表情に基づいて推定された感情、生理的な情報に基づいて推定された感情、身振りなど行為に基づいて推定された感情を示す。これらはいずれもベクトルで表現され、複数の感情値を持つことが可能である。

E_e 関数の五つの構成要素をお互いに独立と仮定すると、 E_e は次のように簡単化される。

$$E_e = \alpha_1 L_e + \alpha_2 S_e + \alpha_3 F_e + \alpha_4 P_e + \alpha_5 A_e \quad (5)$$

α_i ($i=1\sim 5$)は各構成要素の外部刺激に対する重みである。

④ 感情コーパスの構築

言葉遣いに基づく感情推定を行う際には、感情コーパスが必要となる。我々の研究グループではすでに感情コーパス Ren_CECps を構築し、公開している。感情コーパス Ren_CECps とは、感情を表す詳細な言語表現が人手によりタグ付けされた中国語 Weblog 記事の集合である。1,487 文書、11,255 段落、35,096 文、878,164 単語で構成されている。

⑤ ブログからうつ病の検出

言葉遣いに基づく感情推定、音声情報に基づく感情推定、顔表情に基づく感情推定について、いくつかの成果を収めたが、ここでは省略し、文献に譲る。ここに、ブログからうつ病の検出手法を述べる。

Ren_CECps 以外には、日本語の感情コーパスを構築した。日本語については、特に自殺傾向やうつ傾向のあるユーザのブログの収集を行っている。うつ傾向

の有無は、人手によりブログを読み、判断を行う。うつ傾向の判定基準は、著者自身がうつ病であることをブログ上で公言していることである。このようなユーザは、うつ傾向があることを自覚していて、通院経験があったりもする。一方で、うつ傾向であることを自覚していない著者がいたりする。一般に、こうしたユーザ（潜在的うつ傾向ユーザ）のブログを見つけ出すことは難しい。そのため、うつ傾向であることを公言しているユーザを対象に、30名分のブログ（うつ傾向ブログ）と、そうでないユーザ30名分のブログ（非うつ傾向ブログ）とに分類して収集した。計60名の各ユーザごとに、約30日分、述べ1800記事を収集した。

収集したブログの統計データを、表1に示す。

表1：うつ/非うつ傾向ブログコーパスの概要

	うつ	非うつ	合計
記事数	900	900	1,800
形態素総数	273,479	156,082	429,561
形態素異なり数	51,996	38,034	90,030
行数	16,795	13,902	30,697

そして、収集したうつ傾向ブログと非うつ傾向ブログとの間に、何らかの差があるかどうかを調べることにした。本稿では、各ブログにおける感情表現の出現傾向に着目し、特に、時系列的な変化を見ることで、うつ傾向と非うつ傾向のユーザ間の差異について詳しく分析してみる。

まず、うつ傾向ブログまたは非うつ傾向ブログ内に含まれている感情表現の統計をとって見たところ、図1に示す通りの結果となった。このグラフから分かることとして、“hate”、“sorrow”、“fear”、“anger”、“shame”といったネガティブな感情表現は、うつ傾向ブログの方に多く出現しているが、一方で、“joy”、“like”、“relief”のようなポジティブな感情表現も、うつブログの方に多く出現している。“hate”の単語について、圧倒的な差が出ているのは、感情表現辞書中の“hate”の登録語数が最も多いことが原因と考えられる。このことから、全体的な感情表現の出現傾向のみからではうつ傾向の検出が難しいことが分かる。

次に、時系列的な変化の分析について述べる。一般に、ブログユーザは、毎日

ブログを書いている。そのため、毎日の心の動きをとらえるものとしては最適である。収集した各ユーザごとの30日分のブログから導き出した感情表現のポジティブ/ネガティブ出現傾向について解析し、あるユーザについて出現傾向をグラフで表すと、図2のようになった。ポジティブとネガティブの波形が30日という短期間において入り組み複雑に波打っていることが分かる。収集したすべてのブログユーザについて解析を行ったが、短期間の波形の動きを見て、うつかそうでないかを判定することは非常に難しいということが分かった。

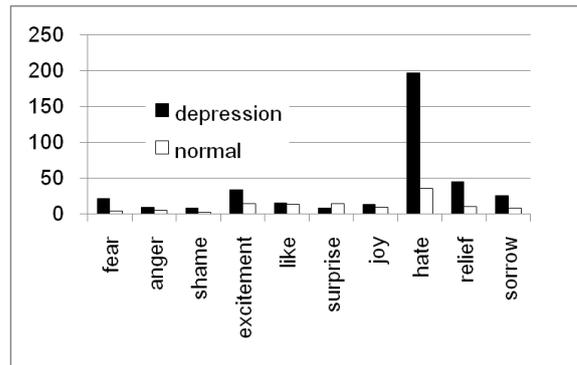


図1：感情表現の出現傾向

図3は、うつ傾向のあるブログユーザについてブログ内のポジティブとネガティブの出現傾向をユーザごとに、可視化したものである。

負の値は、前日からポジティブ/ネガティブの表現がどのように変動したかを示す。図からも分かるように、ネガティブ表現の増加が目立つ結果となった。

ただ、ユーザによっては、うつ傾向があるにも関わらず、ネガティブ表現の減少が大きい場合もあるため、ネガティブ表現の減少の度合いが、すなわちうつ傾向の減少を表すとはいえない。このことから、我々は、うつ傾向の変動をパターン化することを考えた。

まず、うつ傾向の著者が1つのブログ記事内でどの程度の割合で感情表現を用いているかを感情密度と定義し、式(6)によって求める。 $|M|$ は、ブログ中の形態素総数を表し、 pnw_i は、ブログ中のポジティブ/ネガティブ表現を表す。 PN は、ブログ記事中のポジティブ/ネガティブ表現の集合を表す。 w_i は、 pnw_i に対する重みである。この重みは、ブログコーパス中の出現頻度等から算出する。

$$\text{感情密度} = \frac{1}{|M|} \sum_{pnw_i \in PN} w_i \quad (6)$$

この感情密度と、出現する感情表現がポジティブとネガティブのどちらの表現が多いかにより、1つのブログ記事ごとのPN強度を決定する。式(7)にPN強度について示す。

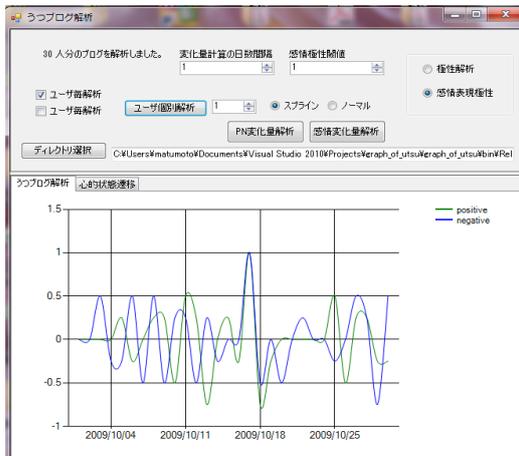


図2：うつ傾向ユーザのP/N変動解析曲線

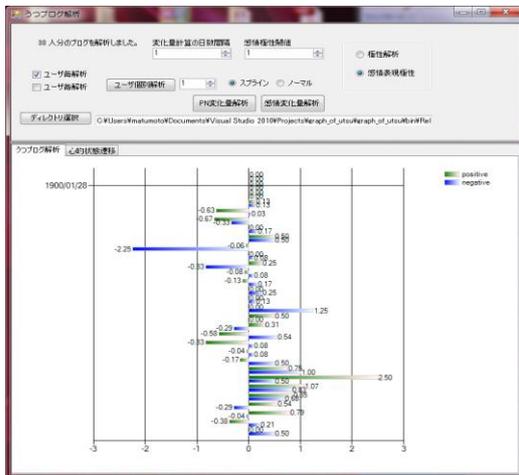


図3：ユーザごとのポジティブ/ネガティブ傾向

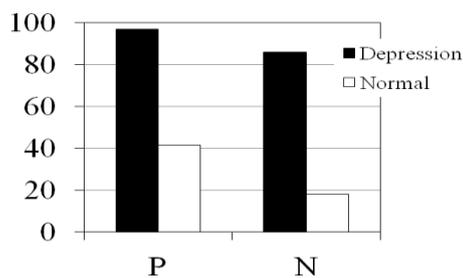


図4：PN強度の平均値比較

$$P\text{強度} = \begin{cases} \text{感情密度} & \text{if } \text{freq}(P) \geq \text{freq}(N) \\ 0 & \text{else} \end{cases}$$

$$N\text{強度} = \begin{cases} \text{感情密度} & \text{if } \text{freq}(N) \geq \text{freq}(P) \\ 0 & \text{else} \end{cases}$$

(7)

PN強度の一日単位での変化を算出し、うつ傾向ブログと非うつ傾向ブログとで、その平均値を比較してみた。図4にその結果を示す。

この結果を見ても分かる通り、うつ傾向がある場合は、うつ傾向が無い場合と比較して、PN強度がP、Nともに高くなる。つまり、うつ傾向の有無をとらえる際に、感情密度の高さが重要になるということである。

また、PN強度の変化を6つの感情変動パターンで表すことにした。ただし、感情密度が0のまま推移するような場合は、以下のどのパターンにも該当しないことにする。

この結果を基に、因子分析を行ったところ、表2に示すような1つの因子が得られた。この因子分析では、因子寄与率が30.6%であった。得られた因子を、本研究では、うつ因子 (Depression Factor) と定義する。また、共通性 (Communality) が特に高いことから、うつ解析に重要となるのが、感情変動パターン PT2、PT6 であることが分かる。PT1 と PT4 は、因子分析の結果でも共通性が低く、PN強度の変動が無いときのパターンであるため、うつ傾向の検出にはそれほど重要ではないと考えられる。

表2：因子分析結果

Pattern	Depression Factor	Communality
PT1	-0.142	0.020
PT2	-0.892	0.795
PT3	0.358	0.128
PT4	-0.165	0.027
PT5	0.448	0.201
PT6	-0.817	0.667

5. 主な発表論文等

[雑誌論文] (計8件)

- ① Xin Kang, Fuji Ren, Predicting Complex Word Emotions and Topics through a Hierarchical Bayesian Network, China Communications, 査読有, Vol. 9, No. 3, 2012, pp. 99-109
http://www.chinacommunications.cn:8080/zgtx/EN/Y2012/V9/I3/99
- ② Peilin Jiang, Fei Wang, Fuji Ren : Semi-Automatic Complex Emotion Categorization and Ontology Construction from Chinese Knowledge,

- China Communications, 査読有, Vol. 9, No. 3, 2012, pp. 28-37
<http://www.chinacommunications.cn:8080/zgtx/EN/Y2012/V9/I3/28>
- ③ Yunong Wu, Kenji Kita, Fuji Ren, Kazuyuki Matsumoto, Xin Kang, Exploring the Importance of Modification Relation for Emotional Keywords Annotation and Emotion Types Recognition, International Journal of Intelligent Engineering and Systems, 査読有, Vol.4, No.4, 2011, pp.19-26
<http://www.inass.org/share/2011123103.pdf>
- ④ Kazuyuki Matsumoto, Yusuke Konishi, Hidemichi Sayama, Fuji Ren, Analysis of Wakamono Kotoba Emotion Corpus and Its Application in Emotion Estimation, International Journal of Advanced Intelligence, 査読有, Vol. 3, No. 1, 2011, pp. 1-24
<http://aia-i.com/ijai/sample/vol3/no1/1-24.pdf>
- ⑤ Kazuyuki Matsumoto, Fuji Ren, Estimation of Word Emotions Based on Part of Speech and Positional Information, Computers in Human Behavior, 査読有, Vol. 27, 2011, pp. 1553-1564
<http://www.sciencedirect.com/science/article/pii/S0747563210003262>
- ⑥ Changqin Quan, Fuji Ren, A blog emotion corpus for emotional expression analysis in Chinese, Computer Speech and Language, 査読有, Vol. 24, No. 1, 2010, pp. 726-749
<http://www.sciencedirect.com/science/article/pii/S0885230810000033>
- ⑦ 三品 賢一, 土屋 誠司, 鈴木 基之, 任福継, コーパスごとの類似度を考慮した用例に基づく感情推定手法の改善, 自然言語処理, 査読有, Vol. 17, No. 4, 2010, pp. 91-110
<http://iss.ndl.go.jp/books/R00000004-I10771999-00>
 [学会発表] (計 18 件)
- ① Liping Mi, Fuji Ren, Hideo Araki, Differences of Japanese Kanji and Kana during the Memory Processing, Proc. of IEEE NLP-KE, Nov. 28, 2011, Tokushima(ホテル千秋閣, 徳島市)
- ② Changqin Quan, Fuji Ren, Selecting clause emotion for sentence emotion recognition, Proc. of IEEE NLP-KE, Nov. 28, 2011, Tokushima(ホテル千秋閣, 徳島市)
- ③ Xin Kang, Fuji Ren, Sampling Latent Emotions and Topics in a Hierarchical Bayesian Network, Proc. of IEEE NLP-KE, Nov. 28, 2011, Tokushima(ホテル千秋閣, 徳島市)
- ④ Fuji Ren, Robotics Cloud and Robotics School, Proc. of IEEE NLP-KE (招待講演), Nov. 28, 2011, Tokushima(ホテル千秋閣, 徳島市)
- ⑤ Ji Li, Fuji Ren, CREATING A CHINESE EMOTION LEXICON BASED ON CORPUS REN-CEPCS, Proc. of IEEE CCIS, Sep. 17, 2011, Beijing(友誼餐館, 北京市)
- ⑥ Changqin Quan, Fuji Ren, An Exploration of Features for Recognizing Word Emotion, Proceedings of the 23rd ICCL, August 24, 2010, Beijing, China
- ⑦ Cheng Wang, Changqin Quan and Fuji Ren, Maximum Entropy Based Emotion Classification of Chinese Blog Sentences, Proceedings of IEEE NLP-KE, August 22, 2010, Beijing, China
- ⑧ Ye Wu and Fuji Ren, Improving Emotion Recognition from Text with Fractionation Training, Proceedings of IEEE NLP-KE, August 22, 2010, Beijing, China
- ⑨ Changqin Quan and Fuji Ren, Construction of a Blog Emotion Corpus for Chinese Emotional Expression Analysis, Proceedings of the 2009 Conference on EMNLP, Aug. 6, 2009, Singapore
6. 研究組織
- (1) 研究代表者
 任 福継 (REN FUJI)
 徳島大学・大学院ソシオテクノサイエンス研究部・教授
 研究者番号: 20264947
- (2) 研究分担者
鈴木 基之 (SUZUKI MOTOYUKI)
 徳島大学・大学院ソシオテクノサイエンス研究部・准教授
 研究者番号: 30282015
 松本 和幸 (MATSUMOTO KAZUYUKI)
 徳島大学・大学院ソシオテクノサイエンス研究部・助教
 研究者番号: 90509754
 谷岡 哲也 (TANIOKA TETSUYA)
 徳島大学・大学院ヘルスバイオサイエンス研究部・教授
 研究者番号: 90319997
- (3) 連携研究者
土屋 誠司 (TSUCHIYA SEIJI)
 同志社大学・理工学部・助教
 研究者番号: 70452654