

音声に基づく気分情報の取得とその提示手法

Sensing of Mental State in Voice and its Expression Method

酒造 正樹^{1*} 荒川 豊² 下地 貴明³ 柴崎 望³

Masaki Shuzo¹, Yutaka Arakawa², Takaaki Shimoji³ and Nozomu Shibasaki³

¹ 神奈川大学 / Kanagawa University

² 奈良先端科学技術大学院大学 / Nara Institute of Science and Technology

³ スマートメディカル株式会社 / Smart Medical Co.,Ltd

Abstract: People under chronic stress cannot have appropriate care without self-reported health status. If their mental stress is quantitatively estimated with small sensors and data mining technologies, we can know their variation of mental state in their daily life. Sharing the information with others will help them not get worse. An application for smart phones which can check the variation of mental state as a life log is discussed in this paper.

1 はじめに

高ストレス社会において、うつ病や何かしらの精神疾患患者が増加している。大学においては、人間関係のもつれや、試験勉強や就職活動への疲れなど様々な要因により、メンタル状態の低調な学生が少なからず存在する。また、筆者の所属する神奈川大学では、2011年の東日本大震災以降、東北被災地に学生ボランティアを派遣している¹が、作業現場において、あるいは活動後に支援疲れを持っている可能性があり、彼らのメンタルケアを行うのが望ましい。同様に、企業においても、職場のストレスチェックが義務化されようとしている。厚生労働省は、2014年の通常国会にて、労働安全衛生法の一部改正を求める予定であり、社会的な法整備が進みつつある²。

本論文においては、メンタルケアを行うにあたり、対象がさほど重症でなくむしろ健康であり、また医師の介在する環境ではない前提のもとにその手法について検討する。まずは自らがメンタル状態を日々把握し、落ち込んだときにはさらに悪化せぬよう、自らエンカレッジできるような仕組みを提案したい。そこで、一検討としてスマートフォンを活用した音声入力に基づく気分情報の取得と、結果の提示手法について報告する。

2 ストレス負荷時の身体反応

一般にストレス負荷時には、声帯の緊張や、心拍の増加、血圧の上昇などが様々な身体反応が起きる(図1)。これらの生理反応は、マイク、心電センサ、血圧センサなどの各種センサを用いて、物理量として観測可能である。これらの物理量(生体情報)を分析することによって、気分状態(狭義に感情情報)の判別が可能になる。複数のセンサにより多角的に解析を行えば、より確度の高い分析結果が得られる。筆者らの研究グループではこの判別技術に関して、様々な研究を行ってきた[1-3]。

本論文においては、用いるセンサ情報は音声に限定する。なぜならば、単一情報による確度低下の懸念があるものの、他の手法に比べセンサ装着の負荷が少なく、また音声はスマートフォンにより簡便に取得可能であり、長期的なライフログとしての用途に向いているからである。

先行研究[4-9]において、音声情報からの気分状態(狭義に感情)の判別は、「怒り」「喜び」「悲しみ」「その他(平常)」などのラベル付きの学習データから抽出した特徴量により判別器を作成し、判別分析を行うものが多い。本論文においては、これらの先行研究における手法を参考に個別の感情を求め、次に情動成分比を計算しこれを気分情報とした。

*連絡先：神奈川大学工学部機械工学科
〒221-8686 神奈川県横浜市神奈川区六角橋 3-27-1
E-mail: shuzo@kanagawa-u.ac.jp

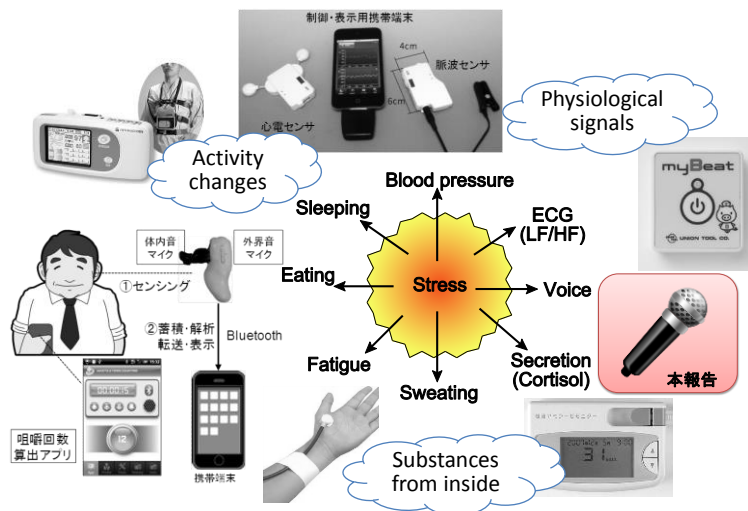


図1 ストレス負荷時の身体への反応とそのモニタリングツール

3 気分情報の取得実験

本章では、軽度の抑うつ傾向の見られる被験者の約1ヶ月の治療期間中の音声の分析により、長期的に気分状態の相対変化が捉えられることを示す。

3.1 目的

各種ストレスのモニタリングツールがある中、最も簡便な音声録音によって、気分の浮き沈みの状態がどの程度把握できるのか、実験により検討を行うことが実験の目的である。

3.2 方法

実験は都内企業勤務の従業員を対象とし、抑うつ傾向が見られる方に同意を取った上で実験協力者とした。抑うつ尺度 (SDS スコア) [10]は、42点～軽度の抑うつ、49点～中等度の抑うつ、57点～重度の抑うつとされており、実験開始前の SDS スコアは、 55.5 ± 9.2 (36～70)であった。症例数は、27名 (男性17, 女性10)であり、平均年齢は、 34 ± 6 歳 (男性23～52歳, 女性26～41歳)であった。これを、医師の治療 (治療期間: 1週～5週)を施した A 群 (14名)と治療のない B 群の2群 (13名)に分け、以下は共通の実験を行った。

実験は2013年3月より開始し、9週間の期間とした。実験期間中、毎晩自身の音声をICレコーダに録音してもらった。就寝前の静音環境下での録音を指示し、発話内容は自由とした。合わせて、参照データとして、週に1回、日本語版 POMS[11]を用いた気分尺度のアンケート調査を実施した。全ての実験終了後に取得した音声ファイルを回収し、気分尺度との相関を分析した。

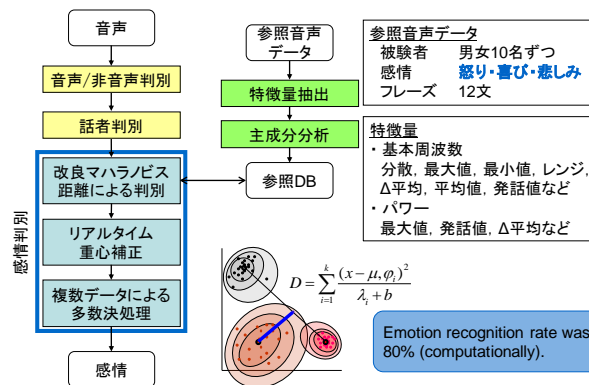


図2 音声情報による感情判別手法の例[7]

3.3 分析

気分尺度は、気分障害度 (Total Mood Disturbance, TMD) 得点に着目した。これは「緊張」「抑うつ」「怒り」「疲労」「混乱」の合計得点から「活気」の得点を引いたものであり、気分障害の大まかな指標として役立つとされている。

音声の解析に関しては、まず、感情判別により、発話ごとに「怒り」「喜び」「悲しみ」「平常」の4状態に分類し、次にポジティブ・ネガティブ情動成分比を計算した。なお、音声から感情判別を行う手法については、先行研究[7] (図2)などに既出である。

3.4 結果

図3に5週までの各群の結果の平均を示した。図3上に示すように、TMDについて、実験中、治療なしの B 群は変化しない。一方で治療ありの A 群の TMD は減少傾向にあり、治療の効果が現れていることがわかる。また、この群に対して、6週目以降、

治療を終了することにより TMD が再び増加傾向にあることが確認された（データは非表示）。

また、音声の解析により得た情動成分について、B 群は TMD と同様に変化しない。一方、A 群については、情動成分が増加する傾向を得た（図 3 下）。

3.5 考察

音声の分析に基づく情動成分の変化の結果がアンケートによる主観評価と良い相関を示したと言える。抑うつ傾向のある実験参加者においては、

- 自分の感情や身体の感覚に気づくことが難しい
 - 感情を表現することが難しい
 - 自己の内面へ眼を向けることが苦手である
- といった失感情症の特徴が見受けられる。つまり、自分が今怒っているのか悲しんでいるのかわからない状態と推測する。治療の効果があがり、抑うつ傾向が改善されるにつれて感情が豊かになり、話す言葉にも情動成分が増加したと思われる。

なお、本論の実験参加者数は 27 名と十分とはいえない。今後は、日常生活環境下でのノイズ処理や個人差の対応などを検討していく必要がある。また、本手法が適合しやすいグループについては再現性を検証する必要がある。

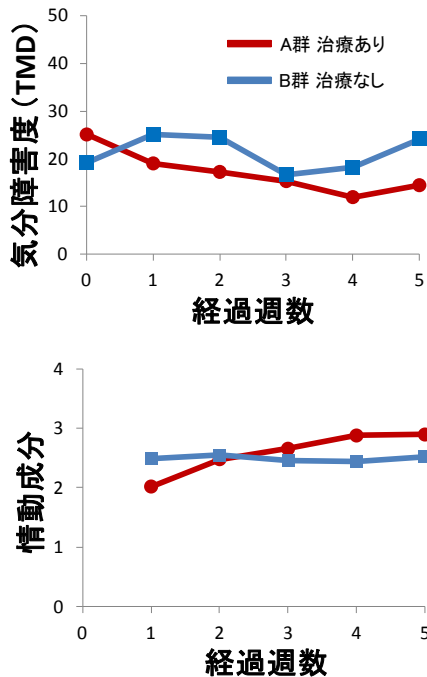


図 3 気分障害度の変化と情動成分の変化

4 情報提示手法の検討

前章までに音声による気分の状態の判別の可能性を示した。提案手法に基づいて、ユーザに対して継続的な精神サポートを提供するためには、モバイルアプリとして提案手法を実装することは重要であると考えられる。一方、その使用手順が煩雑であったり、出力結果がわかりにくいと、提案手法の実際の効果は薄れてしまう可能性がある。そこで、本章ではスマートフォンアプリとして、実装するにあたっての情報入力インターフェースや結果の出力インターフェースに関して検討する。

4.1 音声の入力方法

気分状態の判別を行うためには、理想的にはノイズのない環境下で音声の入力を行う必要がある（図 4 下）。一般に、環境音が多い場所や、複数人が会話をを行うような場所においては誤判別を起こす可能性がある（図 4 上）。そのため現在の実装では、リアルタイムに音声をに入力するのではなく、その場所では写真を撮るだけにとどめ、自宅など静音環境下に移動したあとに、その写真を見ながら回想し、音声を入力する方式を採用した（図 5 左）。この方式は、

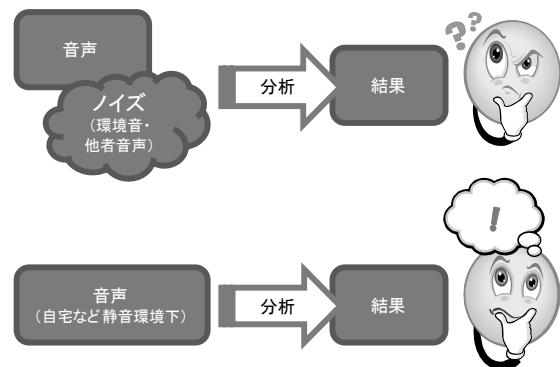


図 4 実環境における音声情報処理

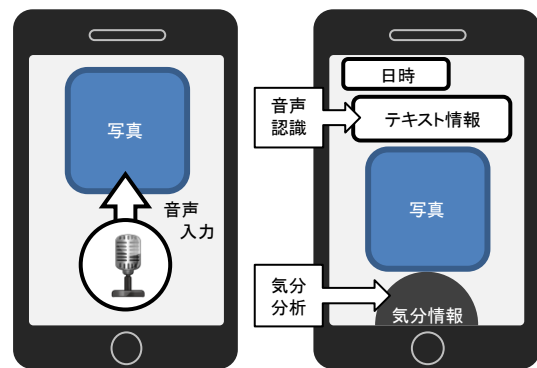


図 5 写真を用いて回想しながら音声入力を行うインターフェース（左）とその結果表示画面（右）

利用者の再現能力に依存してしまうため、将来的には、入力時に適応的なノイズカットを行い、ある程度の雑音であれば、その場で音声を入力できるようにしたいと考えている。ただし、写真を撮ることは、後述するログの表示においても有益であると考えている。

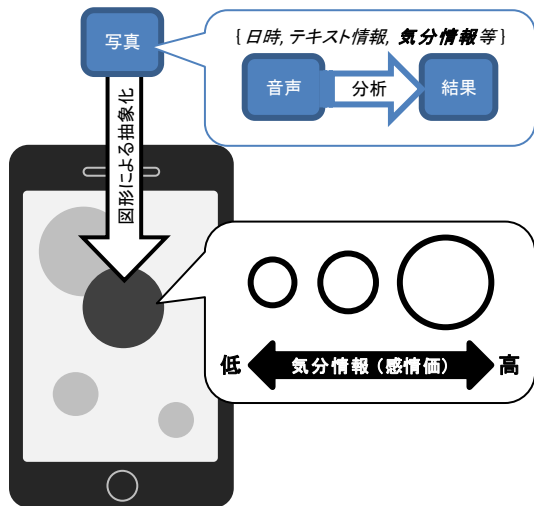
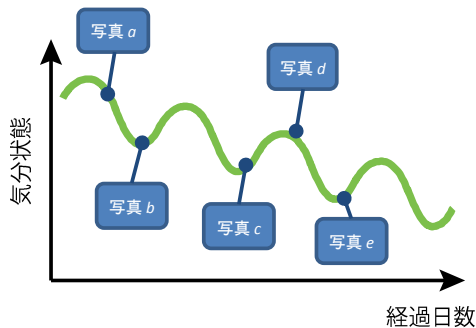


図6 一定期間における気分状態の表示画面



カレンダー表示		ランキング表示	
経過日数	コンテンツ	気分情報 ランキング	コンテンツ
t_1	写真a	1	写真a
t_2	写真b	2	写真d
t_3	写真c	3	写真b
t_4	写真d	4	写真c
t_5	写真e	5	写真e
...

図7 ログの表示

4.2 分析結果の表示方法

気分状態の結果の表示においては、数字や文字などの明示的な情報提示を行わず、図形を用いた抽象的表現を採用し、情動成分比（感情価）の大小に合わせ、図形のサイズが変化するデザインとした。なお、1枚の写真に対する気分状態の結果に対しては、他の期間との相対変化により議論されるものであるから、ユーザには一喜一憂させないことが望ましい。また、並行して走る音声認識ソフトにより、入力したコメントは自動的にテキスト化され表示される。これらの画面配置の例を図5右に示す。

4.3 ログの表示方法

複数のコンテンツ（日時、テキスト情報、気分情報等）を入力した状態での閲覧方法について、一つの画面にはある一定期間の気分状態が抽象的図形の集合として表示される（図6）。なお、個々の図形はランダムに移動している。これに含まれる1つの抽象的図形をクリックすることにより、詳細な結果画面（図5右）が再表示される。結果の詳細表示は、気分を検索のトリガーとして過去の状況を後から回想できることを意味する。ストレスにより、気分が落ち込んだ際に、過去の嬉しかった時など気分が高揚した状況を思い出すことで、落ち込みを和らげることを狙っている。その際、音声だけではなく、写真も提示することによって、回想をしやすくなるため、音声入力前に写真を撮影することは必ずしも手間が増えるだけではないと考えている。

また、この画面では、感情価により画面専有度が大きくなるので、必然的に気分状態の良好であったコンテンツほど回想される機会も多くなる。これも、過去の良好な気分を繰り返し回想することによる効果を狙ったものである。オプションとして気分情報によるランキング表示機能を持つ（図7右下）。

4.4 先行研究との比較

本論文と同じく音声により気分（感情）センシングを行い、また結果の表示も抽象的表現により行われる、類似の先行研究「雰囲気情報端末：障子」[12]と比較する（表1）。据置型の「障子」は、離れた場所に住む家族にネットワークを介して、自らの発話毎の結果をリアルタイムに表示しあい、コミュニケーションを円滑にするものであった。相手の感情などの非明示情報が伝わりにくい電話や遠隔会議システムなどにおいては有用とされる一方で、実環境においては、図4に示す問題点があったほか、主観に基づく“雰囲気”情報を扱ったために、定量的な評価が難しかった。

本論文においては、「障子」の弱点を4.1および4.3に述べた方法で補い、また、さり気なく情報を伝え

表 1 先行研究「障子」との比較

	先行研究	本報告
センシング情報	音声(自動)	音声(手動)
気分(感情)の表示	抽象的表現 (均一サイズ)	抽象的表現 (可変サイズ)
表示場所	遠隔地の据え置き端末	自らのスマートフォン
表示間隔	発話毎, リアルタイム	日毎, 非リアルタイム

るコンセプトを踏襲し、過去のポジティブな体験を自ら省みて、気分の落ち込んだときに自ら励ますことを目標にしている。

5 おわりに

日常のストレスケアに関して、音声の分析により気分の浮き沈みをチェックし、ライフログを取得するスマートフォン向けアプリについて提案を行った。Android OS 向けに実装したアプリのデモを行いながら議論を行いたい。今後は、開発したアプリを用いたユーザスタディを行い、有用性を評価する予定である。

謝辞

愛媛大学医学部の谷川武教授には実験データの解析の際に議論を頂いた。

参考文献

- [1] 井出裕人, 他: 生理指標の多変量解析に基づく個人に依らないストレス推定手法の研究, *情報処理学会第72回全国大会予稿集*, Vol. 2, pp. 561-562 (2010)
- [2] 酒造正樹, 他: ウェアラブル血圧センサの開発, *日本機械学会スポーツ・アンド・ヒューマン・ダイナミクスシンポジウム2011*, pp. 185-189 (2011)
- [3] M. Shuzo, et al.: Wearable Eating Habit Sensing System Using Internal Body Sound, *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, Vol. 4, No. 1, pp. 158-166 (2010)
- [4] 門谷信愛希, 他: 音声に含まれる感情の判別に関する検討, *Technical Report of IEICE, SP*, Vol. 100, No. 522, pp. 43-48 (2000)
- [5] 佐藤秀明, 他: ニューラルネットワークによる感情

音声の分類, *Technical Report of IEICE, NC*, Vol. 101, No. 154, pp. 85-90 (2001)

- [6] 直井克也, 他: 感情に係る特徴量の検討, *Technical Report of IEICE, HIP*, Vol. 105, No. 99, pp. 37-42 (2005)
- [7] 志村誠, 他: 雰囲気コミュニケーション端末における音声を用いた感情抽出手法の研究, *ヒューマンインタフェースシンポジウム2007 論文集*, pp. 593-596 (2007)
- [8] 山本泰史, 他: 自然発話音声を用いた快・不快判別アルゴリズムの提案, *情報処理学会第72回全国大会予稿集*, Vol. 2, pp. 293-294 (2010)
- [9] 酒造正樹, 他: 情動・感情判別のための自然発話音声データベースの構築, *情報処理学会論文誌*, Vol. 52, No. 3, pp. 1185-1194 (2011)
- [10] W.K. Zung: A Self-Rating Depression Scale, *Archives of General Psychiatry*, Vol. 12, pp. 63-70 (1965)
- [11] 赤林朗, 他: POMS(感情プロフィール検査)日本語版の臨床応用の検討, *心身医学*, Vol. 31, No. 7, pp. 577-582 (1991)
- [12] M. Shuzo, et al.: SHOJI: A Communication Terminal for Sending and Receiving Ambient Information, in *Proc. 2009 ASME IDETC/CIE*, pp. 881-888 (2009)

ⁱ KU 東北ボランティア駅伝,
<http://disastersupport.kanagawa-u.ac.jp/>

ⁱⁱ 厚労省, 安衛法改正案の提出へ,
<http://www.mhlw.go.jp/stf/houdou/0000033063.html>