

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 25 日現在

機関番号：32682

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330319

研究課題名(和文) ユーザの操作意図を漏れなく情報機器に伝達することができる音声入力手法の提案

研究課題名(英文) Proposing a novel speech input method that can extract users' operational intentions and convey these for information terminals

研究代表者

小松 孝徳 (KOMATSU, TAKANORI)

明治大学・公私立大学の部局等・准教授

研究者番号：30363716

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：本研究課題では、「ユーザが表現したいけれどもうまく表現できないイメージ」が無意識的に込められているユーザの発したパラ言語情報およびオノマトペを、情報機器に対する入力情報として再利用することができる音声入力手法の提案を行った。具体的には、ユーザがうまく表現できないイメージがパラ言語情報およびオノマトペのどのような要素に反映されているのかを調査し、その結果を解析することでこれらを四次元属性ベクトルとして表現する手法を確立した。そして、ユーザがそれらに込めたイメージを具現化する描画システムを作成し、ユーザの抱いたイメージが適切に反映されているかどうか評価実験を行った。

研究成果の概要(英文)：It is said that user's ambiguous and intuitive feelings are embedded in the onomatopoeic expressions and paralinguistic information in speech sounds. Therefore, an interface system that can use such expressions or information as input information could comprehend such user's feelings, and moreover, this system would contribute to realizing a stress-free speech input method for users. Specifically, I firstly proposed an objective quantification method for onomatopoeic expressions and paralinguistic information, and developed a drawing system that can change the drawing effects of certain target in accordance with the quantified onomatopoeic expressions and paralinguistic information. Based on the positive evaluation for this system, I proposed much more realistic applications especially for the research areas of clinical psychologies.

研究分野：情報学

キーワード：認知科学 感性情報処理 ヒューマンエージェントインタラクション

1. 研究開始当初の背景

近年、インターネットに代表される情報環境の整備は急速に進みつつあり、スマートフォン、タブレット端末などの情報機器を携帯するだけで屋内外問わずにユーザはその恩恵を受けられるようになってきた。また、2011年3月11日に発生した東日本大震災直後の東京都内では、携帯電話や固定電話は完全に不通状態であったものの、インターネット環境はほぼ平常時と変わらずに使用できたことから、その利便性のみならずライフラインとしての頑健性も同時に示されたといえる。よって、このような利便性と頑健性を兼ね備えた情報環境の恩恵を受けることができるユーザ数を増やすことは、情報化社会の実現に向けて喫緊の課題であるといえる。そしてその際には、誰でも情報環境の恩恵を受けられるように、どのような習熟度のユーザであっても扱いやすい情報機器を実現することが必要不可欠となる。現在、習熟度が低いユーザに対しては、限定された機能のみが実装された情報機器を用意することが一般的な対処法であるが(例: 携帯電話に対して通話とメールという基本機能のみを実装したNTTドコモ社の「らくらくホン」は、高齢者層に対して携帯電話を急速に普及させることに成功した)、多種多様な情報のやり取りを可能とする情報環境の利便性に鑑みた場合、使用できる機能を制限するのではなく、ユーザの思い通りに情報環境を利用できるような情報機器を実現することが重要となる。

現在普及している情報機器の多くは、ある特定のチャンネルを通じてのユーザの入力のみを受け付け、その入力に応じた出力をユーザに提示するものである。よって、そのような情報機器に対峙したユーザは、その機器が持つ入力チャンネルに合致した形式にて自らの操作意図を表現し、それを入力する必要がある。その際、情報機器に習熟していないユーザは、自分が情報機器に伝達したい意図は頭の中でイメージできてはいるものの、どのようにそのイメージを情報機器に適した入力情報に変換したらよいか分からない状態にあるといえる。つまり、情報機器に対する習熟度が低いユーザは自らの意図全てを情報機器に入力することができない状態にあるといえ、その状態を克服する情報機器への入力手法の開発は急務であると考えられる。

ユーザが表現したいけれどもうまく表現できないイメージは、ユーザが発するジェスチャ、パラ言語情報(音の高さや大きさといったコトバで表現できない情報)やオノマトペ(擬音語や擬態語などの総称)に無意識的に込められていることがこれまでの研究で報告されている。この知見を踏まえた応募者は、任意のオノマトペを四次元の属性ベクトルとして表現した上でロボットの動作として直観的に表現するシステムを提案し、その有

効性を確認することに成功している。よって情報機器への入力チャンネルが音声情報である場合、この研究の知見を生かして、パラ言語情報を数値化する手法を提案できれば、パラ言語情報やオノマトペとして表現されてしまったユーザが情報機器に伝えることができなかったイメージを入力情報として再利用できるという新たな音声入力手法が実現できると考えた。

2. 研究の目的

そこで本研究課題では、「ユーザが表現したいけれどもうまく表現できないイメージ」が無意識的に込められているユーザの発したパラ言語情報(音の高さや大きさといったコトバで表現できない情報)およびオノマトペ(擬音語や擬態語などの総称)を、情報機器に対する入力情報として再利用することができる音声入力手法の提案を行った。具体的には、ユーザがうまく表現できないイメージを音声にて表出する際に、そのユーザの伝えたいイメージがパラ言語情報のどのような要素に反映されているのかを行動観察(間接観察法)によって調査した。そして、その結果を解析することでパラ言語情報(音の高さや大きさといったコトバで表現できない情報)とオノマトペ(擬音語や擬態語などの総称)とを一括して四次元の数値ベクトルとして表現する手法を検討した。この調査結果を基に、ユーザが表出したパラ言語情報およびオノマトペとしての入力情報を受け付けるアプリケーションとして、それらの情報に込められたイメージを反映するシステムを提案した。そして最終的に、本研究で提案する入力手法が、あらゆる習熟度のユーザであってもストレスなく情報機器に対峙できるかどうかを評価した。

3. 研究の方法

【平成25年度】ユーザがうまく表現できないイメージを音声にて表出する際に、そのユーザの伝えたいイメージがパラ言語情報(音の高さや大きさといったコトバで表現できない情報)のどのような要素に反映されているのかを行動観察(間接観察法)によって調査し、その結果を解析することでパラ言語情報とオノマトペとを一括して四次元の数値ベクトルとして表現する手法を提案した。

【平成26年度以降】ユーザが表出したパラ言語情報およびオノマトペとしての入力情報を受け付けるアプリケーションとして、それらの情報に込められたイメージを反映する描画システムを構築し、そのシステムについての評価実験を行った。

4. 研究成果

【平成25年度】ユーザがうまくコトバとして表現できないイメージを音声にて表出する際に、そのユーザの伝えたいイメージがパ

ラ言語情報（音の高さや大きさといった言葉では表現できないアナログ的な周辺言語情報）のどのような要素に反映しているのかを観察する調査を行った。具体的には、実験参加者に対して目の前の画面に表示されたオノマトペを普通に発話してもらった場合と、そのオノマトペから感じた意図を強調して発話する場合という二種類の音声を探取し、特に、意図を強調した際の発話における音響的な特徴を把握する調査を行った。そして、その音響的な特徴と、オノマトペを表現する四次元ベクトルの各属性値との対応関係について整理した。その結果、キレ属性および躍動感属性が高い値を示しているオノマトペの場合、その意図が強調されると、その発話時間が短くなり、促音が併用されるという傾向が明らかになった。またその一方、柔らかさ属性が高い値を示しているオノマトペの場合、その意図が強調されると、最後の音節にアクセントが付与されたり、促音が使用されるという傾向が明らかとなった。現在のところ、想定していたような時系列的な音響的な特徴は観察されていないために、パラ言語情報の数値化においては、発話時間および促音およびアクセントの検出（つまりパワー情報の変動）に焦点を絞って行うという大まかな方針を立案した。

【平成 26 年度】前年度の研究成果に基づいて、ユーザが発した音声からそのパラ言語情報を抽出し、その意図を数値化して、任意の対象に反映するシステムの開発を行った。具体的なシステムとして、コンピュータ上の描画アプリケーションを作成し、パラ言語情報の数値化が妥当であるかの評価実験を行った。この描画アプリケーションは、画面の中央にボールが表示されており、そのボールに対してオノマトペを音声で入力すると、そのオノマトペおよび音声のパラ言語情報が数値化され、その印象に基づいてボールが動作することで、結果として描画エフェクトを出力するというものである。評価実験の結果、多くのユーザから「自分の思い通りにボールを動かすことができた」「操作がわかりやすかった」といった、高い評価を得ることができた。その一方、現状の問題として、音声から言語情報を抽出する際の精度がそれほど高くないため（それゆえに現状では、入力音声をオノマトペに限定している）、その精度を向上させた上で、オノマトペ以外の言語情報を入力できるようにシステムを拡張することを検討した。

【平成 27 年度】最終年度の研究活動においては、本研究で得られた知見を前年度に開発したシステムよりもさらに具体的なアプリケーションへと昇華するために、臨床心理分野への応用を検討するための網羅的な調査を行った。具体的には、自らの心的状態を表現する際に利用する SUDs と呼ばれる指標と

してオノマトペを利用することで、患者への適切な支援が可能となるという知見が得られたため、患者の音声をリアルタイムで解析するようなシステム案を提案した。

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 5 件)

坂本大介・小松孝徳・五十嵐健夫(2015). パラ言語情報を用いた携帯端末の操作手法, 『ヒューマンインタフェース学会論文誌』, vol.17 (2), 85-96. 査読有.

伊藤惇貴・加納政芳・中村剛士・小松孝徳(2015). オノマトペの音象徴属性値の調整のための一手法, 『人工知能学会論文誌』, vol.30 (1), 364 - 371. 査読有.

岩佐和典・小松孝徳(2015). 視覚による触質感認知と不快感に対する命名の影響 - 触覚オノマトペによる検討 -, 『人工知能学会論文誌』, vol.30 (1), 265 - 273. 査読有.

田中恒彦・岡嶋美代・小松孝徳(2015). 診断横断的行動療法でオノマトペがなぜ有用か?, 『人工知能学会論文誌』, vol.30 (1), 282 - 290. 査読有.

Ito, J., Kanoh, M., Nakamura, T., and Komatsu, T. (2013). Editing Robot Motion Using Phonemic Feature of Onomatopoeia, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol.17 (2), 227-236. 査読有.

〔学会発表〕(計 3 件)

Komatsu, T., and Inaoka, S. (2013). Why are specific onomatopoeias evoked from specific images? - Investigating correlations between image features and quantified onomatopoeias, In *Proceedings of the 2nd Asian Conference on Information Systems (ACIS2013)*, pp. 631-634. 査読有. タイ国・プーケット Merlin Beach Resort 2013/11/1

Sakamoto, D., Komatsu, T., and Igarashi, T. (2013). Voice Augmented Manipulation: Using Paralinguistic Information to Manipulate Mobile Devices, In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2013)*, pp. 69-78. 査読有. ドイツ・ミュンヘン Ludwig - Maximilians - Universität München (LMU) 2013/8/28

Komatsu T., and Kiyokawa, S. (2013). Quantifying Japanese Onomatopoeias Based on Sound Symbolism, In

Proceedings of the 9th International
Symposium on Iconicity in Language and
Literature (Iconicity). 査読無 . 東
京・立教大学 2013/5/3

〔その他〕

ホームページ等

<http://www.tkomat-lab.com>

6 . 研究組織

(1)研究代表者

小松 孝徳 (KOMATSU TAKANORI)
明治大学・総合数理学部・専任准教授
研究者番号：30363716

(2)研究分担者

なし

(3)連携研究者

なし