

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 20 日現在

機関番号：13501

研究種目：挑戦的萌芽研究

研究期間：2012～2013

課題番号：24650350

研究課題名(和文) ライフログを利用した重複肢体不自由者の視線入力型日本語入力システムの研究開発

研究課題名(英文) R&D of the Japanese-Language Input Method using Life Log on an Eye-Controlled Communication Device for Users with Disabilities

研究代表者

小谷 信司 (KOTANI, Shinji)

山梨大学・医学工学総合研究部・教授

研究者番号：80242618

交付決定額(研究期間全体)：(直接経費) 2,900,000円、(間接経費) 870,000円

研究成果の概要(和文)：発話が不可能で両手両足の自由がきかない重複肢体不自由者に対して、視線を利用したコミュニケーション実現を目指している。

過去の研究において、短い語彙入力の場合、有効性が認められたが、長い文章入力の場合で、誤入力が生じると、極端に効率が悪くなることが判明した。そこで、静的属性(知識・経験)、基本属性(時間・空間的情報)、動的属性(周囲・人物情報)を組み合わせ、その状況に応じた予測変換を実現することを目指した。シミュレーション実験と健常者による実験において良好な結果が得られた。現在、支援学校に協力してもらい生徒と一緒に取組を行い、その有効性を検証中である。

研究成果の概要(英文)：We aim to enable the smooth communication of persons physically unable to speak.

In our past study, we proposed three Japanese-language input methods using a portable eye-controlled communication device for users with conditions such as cerebral palsy or amyotrophic lateral sclerosis (ALS). However, these methods require nearly 30 seconds to cycle through one Japanese character. In this research, we suggest a method to estimate the input word using the clues of nearby characters and accumulated experience. In addition, to raise the precision of the prediction, we use the connection between words based on a thesaurus. We have realized precise word conversion via a few input letters, as proved by the result of the simulation experiment.

研究分野：総合領域

科研費の分科・細目：人間医工学・リハビリテーション科学・福祉工学

キーワード：ヒューマンインターフェース パーソナル辞書 ライフログ

1. 研究開始当初の背景

かな漢字変換は、森らにより提案、実装され、現在ではMS-IME、ATOK、Cannaなどが広く利用されている。文脈や知識を利用して少ない入力文字数で正しい変換を行うための様々な研究も行われている。

日本には現在 357 万人の障害者がいて、肢体不自由者は 181 万人である。その中には発話不可能な方が数多くいる。肢体不自由者支援技術は古くから多くの研究がある。その中でも視線入力システムの研究は活発だが、入力の効率性は、ほとんど言及されていない。視線入力を利用した円滑なコミュニケーションを実現できれば、社会的に大きな意味を持つが、日本語視線入力の困難さから、現在まで実現されていない。

2. 研究の目的

山梨大学の補助のもと「視線検出とNIRS（光トポ）による発話不可肢体不自由者コミュニケーションシステムの実現（H21:支援、H22:共生、H23:自立）」を目指している。発話不可肢体不自由者の場合、通常のコミュニケーション手段は全く使えない。3年間の取り組みにより、シンボルの視線選択は可能になったが、日本語を正しく1文字入力するのに30秒近くを要する。健常者ではローマ字入力の場合、3秒程度で1文字の視線入力が可能である。本研究ではライフログ（画像と音声）と現在の周囲状況を手がかりに、時空間的に構成されたパーソナル辞書を切り替えて利用し少ない視線入力文字で的確な日本語変換を実現し、円滑なコミュニケーションを実現することを目的とする。最終的な入力目標時間は健常者と同じ、1文字3秒（現在の1/10）を目指す。

3. 研究の方法

平成24年度は要素研究・基礎研究、平成25年度を応用研究として位置づける。要素研究・基礎研究においては、各項目毎

にインターフェースを決め、未実装な部分は、シミュレーションデータで対応出来る枠組みを作成しておき、全体のスケジュールに支障が生じないようにする。被験者は健常者とするが、障害者のライフログは、平成24年7月から取得する。

応用研究においては、障害者の方に協力を依頼して、システムの検証を行い、継続的改善を行う。障害者の方の場合再現性に欠けるので、実験状況をすべてログとして残し、グラフィカルに表示できる機能を実装する。研究開発環境は、Julius/Julian、および、OpenCVの環境を最大限利用して、開発の効率化を促進する。

将来的な枠組みとして、BMI実現のための基礎データ取得を応用研究時に行う。

システム全体の構成図を図1に示す。

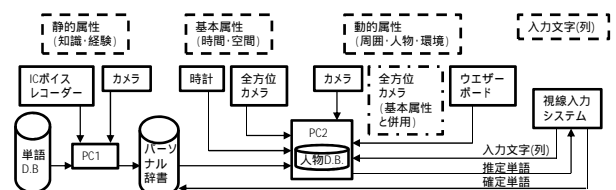


図1：システム全体の構成図

本研究での最大の課題は、一般物体認識である。実環境で行うため、ロバストな手法が必須である。そのため、局所特徴量としてSIFTを用いる。実環境の基本処理ライブラリは、過去の科研費課題で開発済みである。その他にも、文字列取得のためにSVM、各識別のためにB.o.F.、主成分分析、識別器を用いる。

平成25年度は、結合テスト、総合テスト、評価のフェーズである。平成24年度に単体テストの終了しているシステムの結合テストを行う。結合テストにおいては、実環境を想定した学習データでもシステム全体が動作する枠組みを構築する。システム全体の概念図を図2に示す。一点鎖線枠部分は、将来的課題の部分で、現有設備で

ある NIRS (光トポ) と P300 (脳波計) を利用した BMI (ブレインマシンインターフェース) の基礎データ取得のためである。



図 2: システム全体 (将来の BMI 部も含めた) の概念図

定期的に支援学校を訪れ、取得したライフログを本システムに取り込む。作成した視線入力型日本語入力システムを実際に操作してもらい、改良点の洗い出しを行い、問題点を改善する。基本属性、動的属性のデータは、すべてログとして残し、システムの状態、単語の優先順位、変化した属性情報を時系列的にグラフィカルに表示できる機能を実現する。

実験の環境は、教室、食堂、医務室、職員室など、支援学校の教諭、施設の医師、看護師の意見を最大限に取り入れ、多くの他者とのインタラクションが発生するシナリオを複数設定する。

各実験においては、デジタルビデオカメラで実験の様子をすべて撮影する。それぞれの実験終了後、定量的な評価、定性的な評価を行う。特に、支援学校の生徒を対象にした実験では、同一実験を日時をあけて複数回行い、時系列変化に基づく習熟度評価、辞書の更新に基づく学習評価を行う。県内の支援学校に協力を依頼して、被験者を増やし、水平展開する。さらに脳性マヒによる障害だけでなく、重度の ALS の生徒にも被験者依頼し、垂直展開する。

研究の最終年度には、学習会、報告会を開催し、外部評価を受ける。さらに、保護者、支援学校教諭に理解を深めていただく。

4. 研究成果

従来の予測変換技術は、1)過去に入力確定されたものの優先順位を高くする、2)辞書をユーザが指定・選択する、3)文脈を利用する、が一般的である。本研究で対象とするユーザは過去の入力確定データが存在しないため、そのユーザに適切な辞書が存在しない。ユーザに蓄積された情報である静的属性、場所・日時などの基本属性、状況と対象推定による動的属性を融合的に組み合わせ、予測変換精度を向上させることを実現した。表 1 に各項目を示す。

表 1: 各属性の特徴

属性	項目	特徴
静的属性	知識・経験	文字列、累積と忘却
基本属性	時間的情報	単語、累積と忘却
	空間的情報	場所 (屋外では GPS)
動的属性	周囲情報	対象物識別、音声認識、文字識別
	人物情報	人物識別、ライフログ連携

これらの項目が融合的に連携することで、同じ「お」と入力した場合でも、「おはよう」、「おやすみ」、「おめでとう」、「おいしい」、「おおぜい」、「おるごーる」の中から正しい語彙候補の識別を実現した。

静的属性の情報は、カメラによって得られる視覚情報と IC レコーダによって得られる音声情報より生成した。本研究では蓄積された情報をライフログとする。ライフログより、使用頻度の高い固有名詞から構築される語彙データベースを作成する。現在、語彙データベースは手動操作によって作成されているが、自動的に固有名詞を分析する手法を組み合わせることで、より高効率なシステムが期待できる。

基本属性の情報として、時計から日時と現在時刻を取得し、屋外ではGPSから現在位置を取得した。これにより時間と場所に対応した語彙推定が可能となり、挨拶などの基本単語の入力を助ける。

動的属性の情報を得るために、周囲の画像情報、音声情報を取得した。得られた画像には一般物体認識の手法を利用し、周囲にある物体を取得する。肢体不自由者の場合は周囲にある物体を話題にあげることが多いため、この手法によって物体に関する単語を推測する。また、顔認識や人物認識を用いることで人物の特定をし、静的属性との対応をとる。

音声情報はマイクアレイによって取得し、HARKを利用して音源定位を行った。HARKはMUSIC法によって音源分離を行なうので、周囲の雑音に対応したロバストな音声取得が可能である。分離した音声情報にJulius/Julianを用いて語推定を行なった。この語推定により、対話中の人物が話している話題に対応できた。

予測変換を効果的に実装するために、単語ツリーの使用を提案し、実装した。単語ツリーは、シソーラスの考え方をベースにした独自の形式であり、XMLを用いて親和性の高い実装を実現した。使用語彙が比較的少ないと考えられる肢体不自由者の場合、単語ツリーの構築にコストはかからないはずである。ただし、今後健常者が利用することを考えた場合、利用語彙は増えるため、構築のコストも膨大になる。この問題はシソーラスの自動構築を行う手法の利用によって解決が期待できる。

シミュレーション実験を行った結果、単語ツリーを使用して効率の良い予測変換を実現した。複数のワードツリーが活性化した複雑な状況であっても、従来より予測変換よりも効率的な予測変換が可能であった。

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 3 件)

1 「画像センサと距離センサによる 2.5 次元地図の生成と評価」、庄司和晃、渡辺寛望、小谷信司、電気学会論文誌 C、Vol.134, No.2, pp.183-191, 2014, DOI: 10.1541/ieejieiss.134.183、査読有

“R&D of the Japanese Input Method using Life Log on an Eye-Controlled Communication Device for Users with Disabilities”, Shoji Kazuaki, Hiromi Watanabe, Shinji Kotani, Int’l Conf. on PhyCS, Wed8-3, pp.207-213, 2014、査読有

「エピソード平面上に曲線を配置することによるステレオ画像計測」、清水毅、小谷信司、渡辺寛望、丹沢勉、古屋信幸、精密工学会誌、Vol.79, No.4, pp.332-337, 2013、査読有

[学会発表](計 1 件)

「視線入力型日本語入力システムの効率改善に関する研究」、庄司和晃、渡辺寛望、小谷信司、第 31 回日本ロボット学会学術講演会、2H2-06, 20140906, 東京

[その他]

ホームページ等

<http://www.yamanashi.or.jp/~kotani/>

6. 研究組織

(1)研究代表者

小谷 信司 (KOTANI, Shinji)

山梨大学・医学工学総合研究部・教授

研究者番号： 80242618

(2)研究分担者

鈴木 良弥 (SUZUKI, Yoshimi)

山梨大学・医学工学総合研究部・教授

研究者番号： 20206551

渡辺 寛望 (WATANABE, Hiromi)

山梨大学・医学工学総合研究部・助教

研究者番号： 30516943

(3)連携研究者

なし