

論文 / 著書情報
Article / Book Information

題目(和文)	眼球動作入力音声対話支援システムのための高効率入力方式の研究
Title(English)	
著者(和文)	房福明
Author(English)	Fuming Fang
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10546号, 授与年月日:2017年3月26日, 学位の種別:課程博士, 審査員:篠崎 隆宏,小林 隆夫,金子 寛彦,杉野 暢彦,熊澤 逸夫
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10546号, Conferred date:2017/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

専攻： Department of	物理情報システム	専攻	申請学位（専攻分野）： 博士 Academic Degree Requested	（ 工学 ） Doctor of
学生氏名： Student's Name	房 福明		指導教員（主）： Academic Advisor(main)	篠崎 隆宏
			指導教員（副）： Academic Advisor(sub)	

要旨（和文 2000 字程度）

Thesis Summary (approx.2000 Japanese Characters)

疾病や事故等により眼球動作を除く全身の自由運動が失われた場合、発話や筆談のような通常のコミュニケーション手段が全て失われてしまう。このため、残された眼球動作を用いて外部とのコミュニケーションを行うことが必要不可欠となる。このような状況に対応するためのコミュニケーション支援システムとして、これまでにディスプレイ上に文字を表示し眼球動作を利用して文字を選択するソフトウェアキーボード方式や、ディスプレイを用いずに予め決めた眼球動作プロトコルに基づいて文字入力を行うプロトコル入力方式が提案されている。ソフトウェアキーボード方式では1文字あたり3秒程度の比較的速い入力速度が実現できる反面、ディスプレイを用いるため携帯性が低いという問題がある。これに対しプロトコル入力方式はディスプレイが不要であるためポータブルな装置が実現可能という利点がある。プロトコル入力方式としては、これまでに文字の形を眼球の動作でなぞることにより入力を行う手法が提案されている。しかし、従来のシステムでは文字の境界を明確にするため一文字入力するごとに待ち時間が必要であり、一文字あたり8秒と入力に長い時間が必要な欠点がある。

このような背景のもと本研究では眼球動作をもとに音声を経た対話を可能とするため、眼球動作入力音声対話支援システムについて取り組みを行った。プロトタイプとしたシステムでは「上、下、左、右、中」の眼球動作を組合せることで文字を表現し入力プロトコルとした。速い入力速度を得るため、眼球動作に待ち時間は設定せず連続的な入力を行う。眼球動作の検知は眼電位を用いて行い、それぞれの眼球動作を隠れマルコフモデル(HMM)を用いてモデル化することで入力された文字の認識を行う。最後に入力された文字に対応した音声を作成してスピーカから出力する。しかし、眼球動作を連続的に行う場合、眼電位信号は前後の眼球動作の影響を受け特性が変化してしまう問題がある。このため単純に眼球動作ごとにモデル化を行うと、同じモデルに複数の眼電位特性が混ざってしまう、認識精度が非常に低くなってしまふ。また、入力プロトコルは文字ごとに任意に決められた動作の組み合わせであるため、ユーザにとって習得が非常に困難である問題があった。

本研究では眼球動作入力音声対話支援システムにおいて、連続眼球動作の認識精度と入力プロトコルの習得容易性の向上に取り組んだ。まず、認識精度を向上させるために、コンテキスト依存眼動素モデルを提案した。このモデルでは前後眼球動作の影響を考慮して眼電位信号の精密なモデル化を行う。また認識精度をさらに向上させるために、眼電位認識器にN-gram 言語モデルを導入した。N-gram 言語モデルを用いることで、認識処理において言語としての自然さを考慮することが可能となる。健常被験者5人による仮名入力の認識評価を行った場合、眼球動作ごとにモデル化したコンテキスト独立眼動素と比べ、提案するコンテキスト依存眼動素を用いることで仮名認識エラー率を36.0%から1.3%まで大幅に削減できることを示した。N-gram を組み合わせた場合、エラー率はさらに0.9%まで減少し、高精度な連続眼球動作入力が実現できた。

次に、入力プロトコルの習得容易性を向上させるために、入力プロトコルの自動生成に基づいた縮約文字表の提案を行った。この手法はユーザが既に知っている文字表（例えば50音表）を活用し、行と列に動作を指定し組合せることで文字を表現する。さらに、入力に曖昧性を許して隣接する複数の行や列をマージし、1つの動作で指定することで記憶の必要な眼球動作の組み合わせを削減し、習得容易性を向上させる。入力した動作列が複数の文字列に対応する曖昧性が生じるが、N-gram を利用することでもっとも自然な文章を出力する。プロトコルの自動生成では習得容易性や入力曖昧性などを考慮して最適なマージ方法を探索する。被験者評価実験では、マージを行わなかった50音表をベースライン入力プロトコルとして、縮約した50音表と比較した。被験者には、できるだけ正しく入力できるように入力プロトコルを学習することを求めた。この実験においてベースライン入力プロトコルの学習には5.1分かかるとのに対して、縮約文字表では2.4分しかかからないことを示した。提案法を実装した眼球動作入力音声対話支援システムを用いた場合の入力速度は一文字あたり1.9秒となり、従来の入力方式に基づいたシステムより大幅に速い速度を実現した。

備考：論文要旨は、和文2000字と英文300語を1部ずつ提出するか、もしくは英文800語を1部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1 copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).

(博士課程)
Doctoral Program

論文要旨

THESIS SUMMARY

専攻： 物理情報システム 専攻
Department of
学生氏名： 房 福明
Student's Name

申請学位 (専攻分野)： 博士 (工学)
Academic Degree Requested Doctor of
指導教員 (主)： 篠崎 隆宏
Academic Advisor(main)
指導教員 (副)：
Academic Advisor(sub)

要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words)

Human-computer interface systems of which input is based on eye movements can serve as a means of communication for patients with locked-in syndrome. There are various such systems developed. However, some of these systems suffer from non-portability for lack of a computer screen, and others without a screen suffer from slow input speed because they require a pause between input characters to simplify the automatic recognition process. This study has previously developed a non-screen-based continuous eye-input system, in which a pause is not required between characters and fast input speed can be achieved. In this system, the characters were encoded by combining a set of eye motions, and an input protocol was used to represent the coding. Biomedical signal electrooculography (EOG) was used to detect eye motions and hidden Markov model (HMM) was used to recognize the input characters from EOG. However, since the continuous eye movements were fused and influence each other, accurate recognition was very difficult. Further, since the input protocol was very complex, it was very hard for the users to learn. In order to achieve the continuous eye-input system, this study proposes a context-dependent EOG model to improve the recognition performance and an automatic input protocol recommendation method to enhance learnability of the input protocol. With the context-dependent EOG model, the influences of the preceding and succeeding eye movements are considered and EOG characteristics are precisely distinguished and accurate recognition is achieved. For the automatic input protocol recommendation method, a language table that the users already known is used. Then each of the merged adjacent rows/columns is associated to an eye motions to simplify the table. Because the input motions correspond to multiple character sequences, an N-gram is used to extract the most natural character sequence. Experiment showed a faster input speed than the existing systems without a screen, higher recognition performance than the previous modeling method that does not consider influence of the contexts, and more user-friendly input protocol than the input protocol used in the previously continuous eye-input system.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).