

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)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	房 福明		
		氏 名	職 名		氏 名	職 名
論文審査 審査員	主査	篠崎 隆宏	准教授	審査員	杉野 暢彦	准教授
	審査員	小林 隆夫	教授			
		金子 寛彦	准教授			
		熊澤 逸夫	教授			

論文審査の要旨 (2000 字程度)

本論文は「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」と題し、全6章より構成されている。

第1章「序論」では、研究の社会的位置づけについて説明するとともに当該分野における過去の研究を概観し課題を挙げて本研究の目的を述べている。具体的には、本論文は筋委縮性側索硬化症 (ALS) 等により眼球以外の動作が失われた患者に対して、眼球動作をもとに発音を指定し合成音声を出力することで音声コミュニケーションを支援するシステムの実現に取り組むと述べている。従来の眼球動作入力手法を大別すると、ディスプレイを用いる方式と用いない方式に分類できる。ディスプレイを用いない手法は携帯性に優れる反面、システム側の眼球動作認識性能の制約から一文字ずつ時間間隔をあけて文字情報を入力する必要があり、入力速度が遅い問題がある。このような背景のもと、本論文における研究では、眼電信号にもとづく高精度な連続眼球動作認識システムを実現すること、および入力速度が速くかつ利用者にとって習得が容易な入力方式を実現することを目的とすると述べている。

第2章「統計的パターン認識」では、連続眼球動作認識システムを実現する上での基礎となる、時系列信号を対象とした一般的なパターン認識技術について解説を行っている。まず基本的な枠組みの定式化として、雑音伝送路モデルについて説明を行っている。ついで統計的な認識システムの実現方法として隠れマルコフモデル (HMM)、N-gram モデル、重み付き有限状態トランスデューサ (WFST)、および探索アルゴリズムについて説明を行っている。

第3章「眼球動作入力音声対話支援システム」では、本研究においてベースとして用いている眼球動作入力音声対話支援システムのフレームワークについて、説明を行っている。そしてそのフレームワークをもとに取り組みを行った、連続的な眼球動作入力が可能で文字情報の入力に待ち時間が不要となる、HMM を用いた眼球動作認識システムについて説明を行っている。そのシステムでは認識性能が不十分であり、また入力方式が複雑で利用者が使用方法を習得するのが非常に困難である問題点が存在することが、実験結果の数値を示しながら述べられている。

第4章「認識精度の向上」では、待ち時間を設けることなく連続的に文字情報入力を行う際に必要となるカテゴリカルな眼球動作の高精度な認識を実現する方法として、コンテキスト依存眼動素 HMM モデルの提案を行っている。これは、認識対象となる眼球動作のモデル化において前後の眼球動作の影響をモデル化するために、前後の眼球動作に応じて場合分けをして眼球動作のモデル化を行うものである。これにより前後の眼球動作の影響を明示的にモデル化することが可能となる。しかしこの方法を単純に適用した場合、モデルのパラメタ数が眼球動作のカテゴリ数の3乗に比例して大幅に増えてしまい、有限データからのモデル学習が難しくなる問題がある。そこで、木構造クラスタリングを応用した状態共有モデルについて提案を行っている。また、さらに認識性能を向上させるために、N-gram モデルを組み合わせている。WFST を用いて認識システムを構築し、認識実験を行っている。そして、認識性能が従来法と比べ大幅に改善することを示している。

第5章「入力プロトコルの覚えやすさと速度の向上」では、仮名文字入力の方式として、50音表の文字配列を基にした方式と、文字の形を直接眼球の動作でなぞる連続画なぞり方式について提案している。50音表を用いる方式では、表の複数の行や列に同じ眼球動作を対応させることで入力速度や習得性をさらに向上させる縮約文字表方式についても提案を行っている。画なぞり方式については類似の従来法が存在するが、連続的な入力が可能で文字ごとに区切りのための待ち時間を設ける必要がない点が新規であると述べている。評価実験により、縮約文字表方式や連続画なぞり方式を用いることで入力速度や習得容易性が向上することが示されている。また縮約文字表方式は入力速度の点で、連続画なぞり方式は習得容易性の点で、それぞれ特に優れていることが示されている。

第6章では、本研究で得られた成果をまとめている。

以上を要するに、本論文では眼球動作をもとにした音声コミュニケーション支援システムとして高速な入力が可能で利用者にとって習得が容易な利点を有する2種類の連続入力方式を提案するとともに、そこで必要となる高精度な連続眼球動作認識システムを実現している。これらの成果は従来システムの問題点を大きく改善するものであり、眼球動作を基にした実用的なコミュニケーション支援システムを実現する上で工学上貢献するところが大きい。また福祉面での貢献も期待されるものである。よって本論文は博士（工学）の学位論文として価値があるものと認められる。