

令和 元年 6月14日現在

機関番号：24403

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K21287

研究課題名(和文)聴覚障がい者とのストレスフリーな対話に向けたフィンガースペリングの新認識システム

研究課題名(英文)A New Recognition System of Finger-spelled Sign Language Toward Stress-free Communication with Hearing-impaired People

研究代表者

井上 勝文(Inoue, Katsufumi)

大阪府立大学・工学(系)研究科(研究院)・助教

研究者番号：50733804

交付決定額(研究期間全体):(直接経費) 3,000,000円

研究成果の概要(和文):本研究では,聴覚障がい者とのストレスフリーなコミュニケーションの実現に向けて,日本語のフィンガースペリング(以後,指文字と表記)の新認識システムを開発した。

具体的には,日本語の指文字に含まれる静止文字の認識精度向上と連続表現された指文字のスポッティングに取り組んだ。実験結果より,静止文字を89.9%で認識できた。また,18単語のスポッティングの実験において,平均で60.4%の文字分割成功率を達成した。さらに,これらの技術を駆使し,日本語の指文字を用いたタイピングシステムのプロトタイプを構築した。ユーザアンケートより,他の手段に対する代替手段として有効であることを確認した。

研究成果の学術的意義や社会的意義

日本語のフィンガースペリング(指文字)において,これまでの研究では,指文字に含まれる静止文字を一文字ずつ認識することに焦点が当てられていた。これに対し本研究では,この議論から一歩踏み出し,日本語の指文字における動作文字の認識と,さらに連続表現された指文字を一文字一文字どの区間で表現されていたかを解析するスポッティングに取り組んだ。実験の結果,平均で60.4%の文字分割成功率を達成し,静止文字一文字一文字の認識という話で止まっていた当該分野に新たな議論の種をまくことができたと考える。また指文字によるタイピングシステムを試作し,アンケート調査より筆談等の代替手段となれる可能性があることを示した。

研究成果の概要(英文):In this research, we proposed a new recognition system of Japanese finger-spelled sign language (JFSL) toward stress-free communication with hearing-impaired people. The detail is as follows.

First, we improve the recognition performance of static signs in JFSL with OpenPose which is a finger pose estimation method. From the experimental results, we achieved 89.9% of average recognition accuracy. Second, we proposed a spotting system of sequential representation of JFSL. From the experimental results with 18 words, we achieved 60.4% of average division success rate for spotting. Finally, we constructed a prototype of typing system with JFSL by using the proposed spotting method mentioned above. This system can convert the Hiragana representation into its Kanji ones. From the questionnaire with 5 subject, we confirmed that this system has a potential for replacing the existing tools such as keyboard and hand writing.

研究分野：知覚情報処理

キーワード：フィンガースペリング認識 RGB-Dカメラ スポッティング 距離情報 ジェスチャー認識 動作解析

1. 研究開始当初の背景

日本が抱える社会問題の一つに聴覚障がい者の問題がある。厚生労働省の調査 (<http://www.mhlw.go.jp/toukei/saikin/hw/shintai/06/>)によると、聴覚障がい者の数は30万人以上と言われており、また、高齢化社会でその人数は増加傾向にある。聴覚障がい者とのコミュニケーション手段の一つに手話があるが、手話通訳士として登録されている人は約3千人と、聴覚障がい者の人数と比較すると圧倒的に少ないという現状がある。また手話は、その語彙数が数千個以上存在し、日常語だけでも数百個以上あると言われていたため、高齢者だけでなく、健常者でも習得するのが難しいという問題がある。これに対し、視覚言語(手話)の一種でフィンガースペリング(以後、指文字と表記する)と呼ばれる書記言語の綴を一文字一文字指の形で表現するものがある。日本語の指文字を図1に示す。日本語の指文字は、その表現方法において静止文字と動作文字の二つに分類できる。前者は、清音のうち「の」、「も」、「り」、「を」、「ん」の5文字を除く41文字が該当し、後者は残りの文字全てである。このように、日本語の指文字は、高々ひらがな50音(濁音、半濁音、小書き文字を含む)分を覚えれば使用でき、手話単語にない単語を表現する際に一般的に用いられている。指文字は、ひらがなのみならずアルファベット等も存在するため、これを高精度で認識できれば、聴覚障がい者と円滑なコミュニケーションが実現できると考え、世界中で研究されている。しかし、既存のシステムでは認識精度が低く、実用にまで至っていないため、ユーザがシステムの使用にストレスを感じている問題がある。

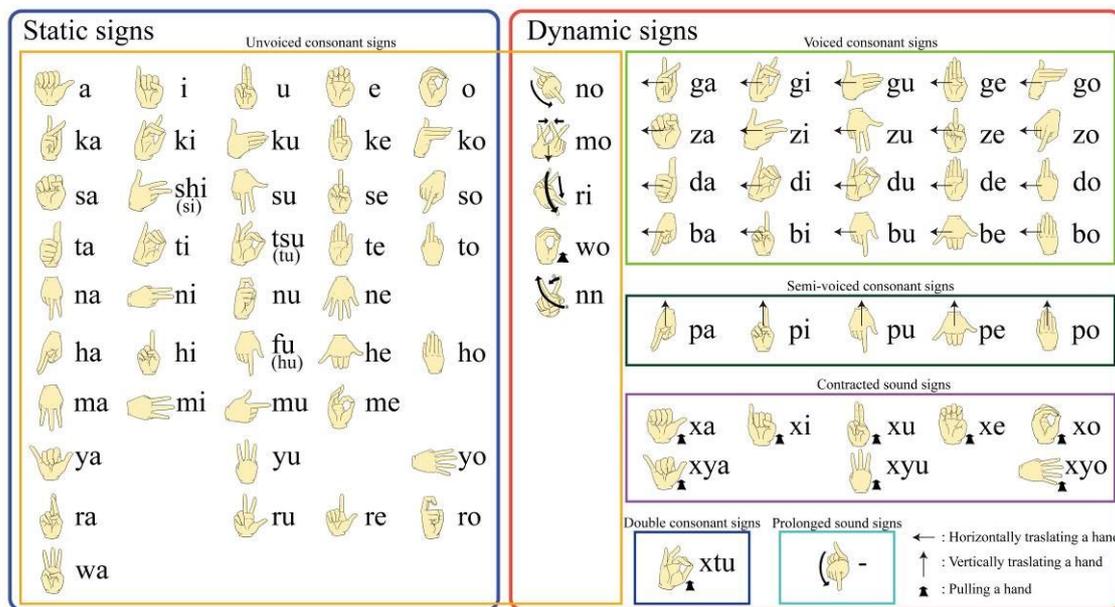


図1 日本語の指文字

指文字認識で最も研究されているアプローチは、カメラで話者を撮影した画像より手領域を抽出し、その手や指の形を認識するものである。このアプローチでは、画像からの手領域の抽出精度と抽出した領域の計上の認識精度向上が大きな課題であり、世界中で多数の手法が提案されている。申請者もこれまでに日本語の指文字認識において、対象物までの距離も測定できるRGB-Dカメラを用いた手領域抽出手法や、HOGといった画像特徴量を用いた静止文字認識手法、動作方向を考慮したルールベースの動作文字認識手法を提案してきた。しかし申請者も含めこれまでの日本語の指文字認識に関する研究では、静止文字の認識を主に対象としており、動作文字の認識、表現する文字を連続で切り替えた際の各文字の時系列的な切り出し、文字の連続性を考慮した単語認識、といった課題を扱う研究はほとんどなかった。

2. 研究の目的

日本語の指文字において、大部分の動作文字は、ベースとなる静止文字を表現しつつ手を動かして表現するため、静止文字の認識精度向上は、指文字認識システムの根幹を成す課題である。このため、これまで多くの研究が成されてきた。しかし、静止文字のみでは聴覚障がい者と円滑なコミュニケーションをとることはできないため、申請者は静止文字認識の問題のみに固執せずに、上述の残された課題について議論の駒を進める。

そこで本研究では、連続表現された日本語の指文字を一文字一文字表現された区間を特定(スポッティング)しつつ、各文字を高精度に認識するシステムの確立である。また、この認識結果から連続表現されたひらがな列を漢字に変換するタイピングシステムを実現する。これによりさらなる聴覚障がい者とのコミュニケーションの促進と、システムの誤認識による会話ストレスを軽減させることを本研究の目的とする。

3. 研究の方法

本研究では、上述の目的を実現すべく、以下の三種類の研究を行った。

(1) 静止文字認識精度向上

上述の通り、日本語の指文字において、静止文字の認識精度向上はシステムの根幹を成す課題であるため、研究期間を通して取り組んだ。具体的には、画像特徴量と距離情報の組み合わせ、および指骨格の形状特徴量を用いて静止文字の認識精度向上を図った。以下で詳細を述べる。

(1.1) 画像特徴量と距離情報の組み合わせによる静止文字認識精度向上

本研究課題を開始する前に、申請者は RGB-D カメラから得られる手領域画像から HOG 特徴量と呼ばれる画像特徴量を用いた静止文字認識手法を提案していた。そこで、まずこの既存手法を改良した。RGB-D カメラは、画像のみならず距離情報も取得できるため、手領域の凹凸情報なども利用した。本研究では、文献[a]を参考に、手領域の凹凸情報等を抽出し、画像特徴量と組み合わせた特徴量を SVM(サポートベクトルマシン)と呼ばれる機械学習の手法を用いて静止文字の認識精度を改善した。

[a] 井上, 小山, 齋藤, “距離画像を用いた単一指文字認識”, 電子情報通信学会技術研究報告, PRMU2013-167, pp. 195-199, 2013.

(1.2) 指骨格の形状特徴量を用いた静止文字認識精度向上

OpenPose[b]と呼ばれる深層学習技術によって手骨格情報を推定する手法を用いた静止文字認識手法を提案した。具体的には、OpenPose を用いて指の関節点位置と手首点の位置を推定し、手首点から見た各関節点の相対位置情報を特徴量とし、SVM を用いて静止文字の認識精度を向上させた。

(2) スポットティング手法

連続表現された指文字から単語を認識するためには、まず指文字一文字一文字が表現された区間を特定しつつ、その文字が何かを認識する必要がある。本研究では、SOM(自己組織化マップ)を用いた手法と、SVM とルールベースの手法の二つを提案した。以下で詳細を述べる。

(2.1) SOM を用いたスポットティング手法

SOM は $W \times W$ のマップで構成され、マップを構成する各ニューロンは参照ベクトルと指文字のラベルを持つ。参照ベクトルは SOM への入力ベクトルと同じ次元数で構成される。本研究では SOM への入力として、静止文字と動作文字の動き始めと動き終わりの手領域から得られる HOG 特徴量を用いる。そして、これらを用いて参照ベクトルを予め学習させておく。テスト時に、入力ベクトルと参照ベクトルとの距離が最小となるニューロンを求め、これに登録されている指文字のラベルを参照する。また、この計算と同時に手の移動量を求め、移動量が大きい場合動作文字が表現されたと判断し、小さければ静止文字と判断する。上述のラベルと手の移動方向等の情報から静止文字及び動作文字を認識する。

また、表現されている区間として、静止文字の場合同じ文字が表現されている区間をその文字が表現されていた区間とし、動作文字の場合は、移動量から動作開始と終了タイミングを取得して、その区間を表現区間とし、指文字のスポットティングを実現した。

(2.2) SVM を用いたスポットティング手法

SOM を用いたスポットティング手法では、SOM 自体の指文字認識精度が低かったため、同じ文字を表現しているにもかかわらず安定して認識できないという問題が生じた。そこで、SOM よりも識別性能の高い SVM を SOM と置き換えることで精度向上を図った。

(3) 指文字を用いたタイピングシステムのプロトタイプ構築

上述の SVM を用いたスポットティング手法を応用し、連続表現された指文字のひらがな列を漢字変換するタイピングシステムのプロトタイプを構築した。本プロトタイプのタイピングシステムでは、MeCab に基づく漢字変換エンジン mecab-skkserv を使用した。ユーザに指文字を連続で表現してもらい、スポットティングして得られたひらがな列を mecab-skkserv で漢字予測変換された第一候補を提示し、本プロトタイプのユーザビリティ性を調査した。

4. 研究成果

(1) 静止文字認識精度

画像特徴量と距離情報を組み合わせた手法を評価するために、8 名の実験参加者に各静止文字を 10 回ずつ表現してもらい撮影したデータを用いた。7 名分のデータで学習し 1 名分をテストするというルーチンを 8 回行い、静止文字の平均認識精度を求めた。本実験では、41 種の静止文字を 88.4%で認識できることを確認した。

また、指骨格の形状特徴量を用いた手法の評価実験では、18 名の実験参加者に各静止文字を 10 回ずつ表現してもらい撮影下データを用いた。17 名分のデータで学習し、1 名分をテストするというルーチンを 18 回行い、静止文字の平均認識精度を求めた。本実験では、41 種の静止文字を 89.9%で認識できることを確認した。

(2) スポットティング手法

SOM を用いたスポットティング手法を評価するために 8 名の実験参加者に静止文字および動作文字の動作開始時と動作終了時の指も形を 10 回ずつ表現してもらい SOM を学習した。学習で得られた SOM を図 2 に示す。また、5 名の実験参加者に、「あした(明日)」、「きのう(昨日)」、「えいご(英語)」、「れしび(レシピ)」、「いしゃ(医者)」、「がが(画家)」、「のり(海苔)」の 7 つの単語を 10 回ずつ表現してもらった。評価方法としては、表現される語句を構成する文字数と同じ指文字数が検出できた割合を示す文字分割成功率と、各語句を構成する文字と認識された文字

が全て一致する割合を示す単語認識成功率の二つの指標を用いた。実験の結果、平均文字分割成功率 66.4%、平均単語認識成功率 52.1%を達成した。

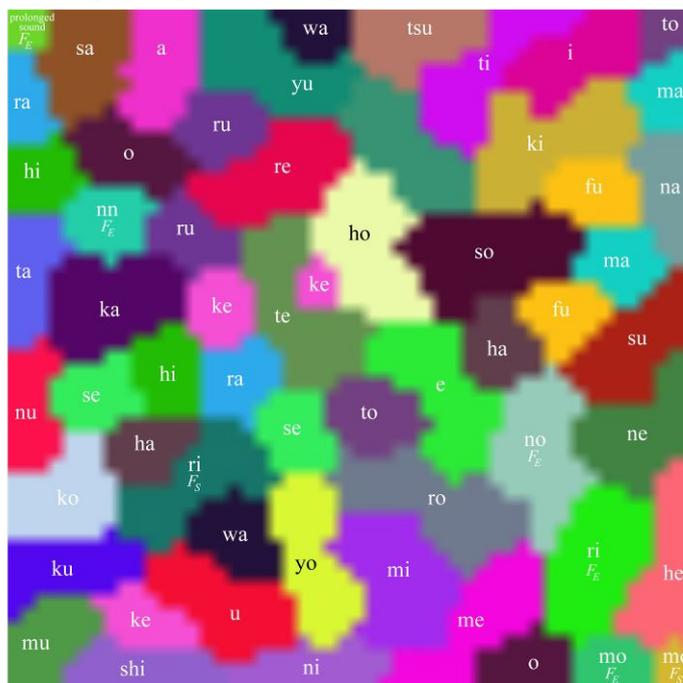


図 2 SOM の学習結果

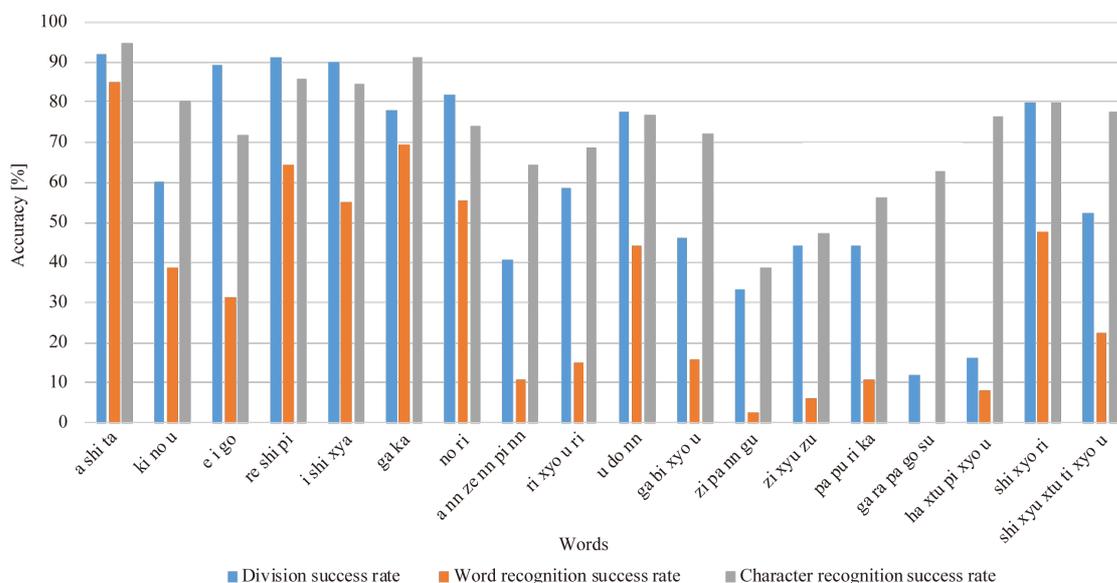


図 3 18 単語のスポッティング結果

また、SVM を用いたスポッティング手法を評価するために、SOM を用いた手法の評価で文字射た 7 単語に、「あんぜんぴん(安全ピン)」、「りょうり(料理)」、「うどん」、「がびょう(画鋸)」、「じばんぐ(ジパンク)」、「じゅず(数珠)」、「ぱぶりか(パプリカ)」、「がらぱごす(ガラパゴス)」、「はっぴょう(発表)」、「しより(処理)」、「しゅっちょう(出張)」の単語を加えた、計 18 単語を 16 名の実験参加者に 10 回ずつ表現してもらい精度を評価した。本実験では、SOM で用いた二つの評価指標に、表現された文字毎の認識精度を表す文字認識成功率を追加して評価した。結果として、図 3 に示すように、平均で文字分割成功率 60.4%、単語認識成功率 32.5%、文字認識成功率 72.4%を達成した。

(3) 指文字によるプロトタイプタイピングシステム

SVM に基づくスポッティング手法と mecab-skkserv を用いた指文字によりタイピングシステムのプロトタイプを作成した。図 4 にプロトタイプシステムの処理の概要を、図 5 にプロトタイプシステムの動作例を示す。図 5 に示すように、スポッティング処理で得られたひらがな列を漢字変換して提示している。また、本プロトタイプシステムのユーザビリティを評価するため、「Q1: スポッティング処理がスムーズである。」、「Q2: 指文字表現者側の制約条件(システムからの距離、表現角度や姿勢)が優しい。」、「Q3: 聴覚障がい者とストレスなくコミュニケーションを行うことができると思う。」、「Q4: 本システムは筆談やキーボードの他の手段として有効

である」という4つのアンケートを実験参加者5名に対して行った。結果を図6に示す。図6より、システムの処理精度や円滑な意思疎通の実現性についてまだ否定的な意見が多く、システムの改善が必要であると考えられる。一方、システムの制約条件や他の手法の代替手段として、肯定的な意見が多かったため、システムの柔軟性や必要性を確認することができた。このことから、プロトタイプとして一定のユーザビリティ性を有していることを確認した。

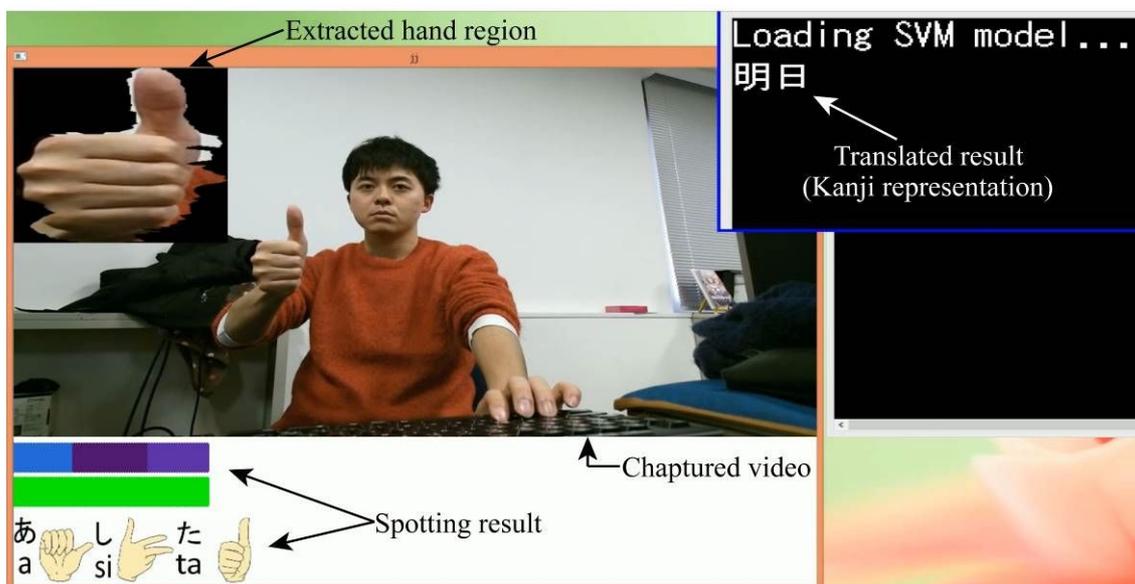


図5 指文字によるタイピングシステムの動作例

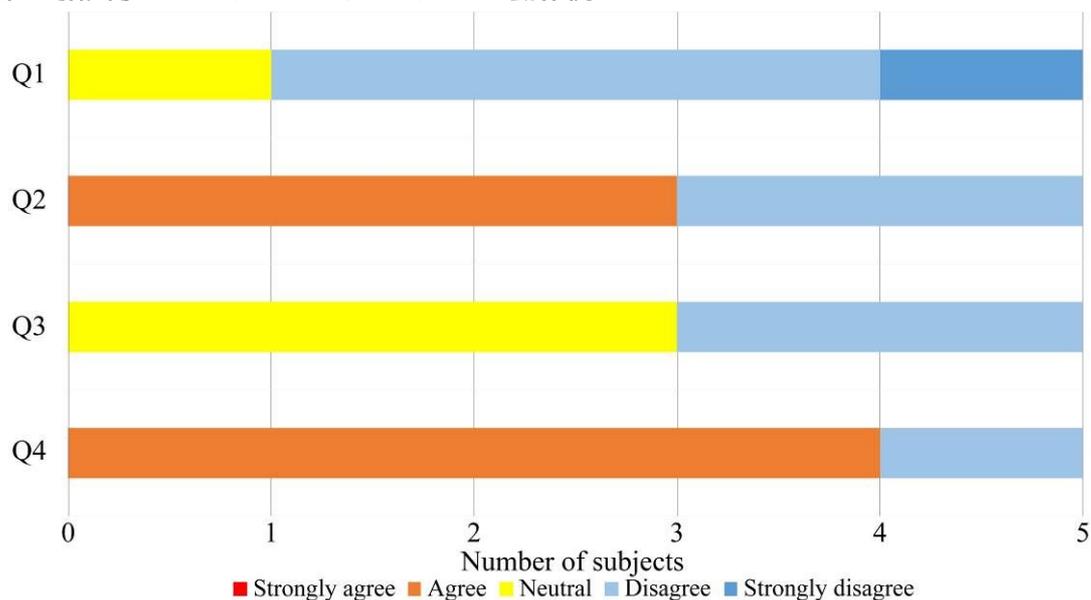


図6 ユーザアンケート結果

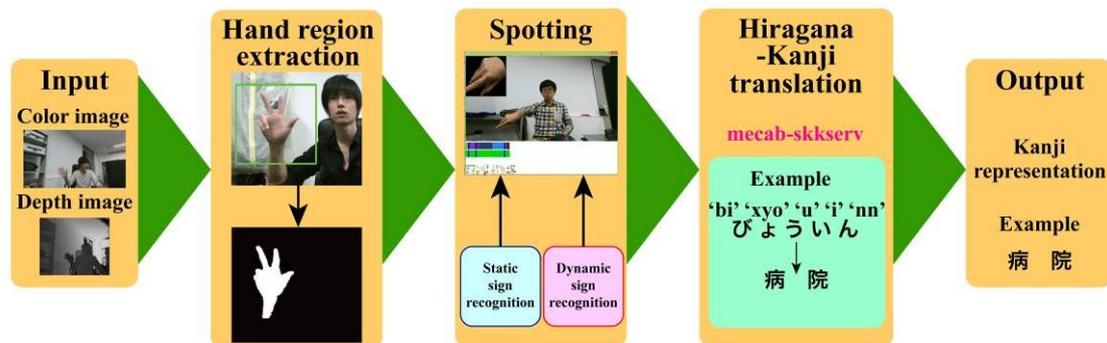


図 4 指文字によるタイピングシステムの処理の流れ

5. 主な発表論文等

〔雑誌論文〕(計 1 件)

1. 井上勝文, 白石孝弥, 松岡遼, 吉岡理文, “ RGB-D カメラに基づく連続表現された日本語の指文字のスポットティング解析 ”, 電気学会論文誌 C(電子・情報・システム部門誌), 138, 10, pp.1230-1241 (2018-10), 査読あり

〔学会発表〕(計 8 件)

1. 松岡遼, 井上勝文, 吉岡理文, “ 日本語の指文字を用いたタイピングシステムの検討 ”, 平成 30 年度計測自動制御学会関西支部・システム制御情報学会若手研究発表会, B3-4, pp.121-124 (2019-1).

2. 井上勝文, “ RGB-D カメラを用いた日本語の指文字認識 ”, 平成 30 年電気関係学会関西連合大会, G12-8, 2 pages (2018-12).

3. 丸山瑞己, 奥村司, 松岡遼, 井上勝文, 吉岡理文, “ OpenPose を用いた日本語の指文字における静止文字認識 ”, 第 5 回サイレント音声認識ワークショップ, 1 page (2018-9).

4. 井上勝文, 丸山瑞己, 奥村司, 松岡遼, 吉岡理文, “ OpenPose を用いた日本語の指文字における静止文字認識に関する検討 ”, 第 21 回画像の認識・理解シンポジウム, PS1-23, 2 pages (2018-8).

5. 松岡遼, 井上勝文, 吉岡理文, “ RGB-D カメラを用いた日本語の指文字における静止文字認識 ”, 平成 29 年電気学会電子・情報・システム部門大会, TC20-1, pp.766-771 (2017-9).

6. 松岡遼, 白石孝弥, 井上勝文, 吉岡理文, “ 日本語の指文字における単語認識に向けたスポットティング解析 ”, 第 4 回サイレント音声認識ワークショップ, 1 page (2017-9).

7. 井上勝文, 白石孝弥, 吉岡理文, “ Kinect を用いた日本語の指文字のスポットティング解析 ”, 第 20 回画像の認識・理解シンポジウム, PS3-54, 2 pages (2017-8).

8. 白石孝弥, 井上勝文, 吉岡理文, “ 距離画像センサに基づく SOM を用いた日本語の指文字認識とスポットティング手法 ”, 第 19 回画像の認識・理解シンポジウム, PS1-26, 2 pages (2016-8).

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ等

6. 研究組織

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。