

3Q-4 漢字字画による日本語漢字入力補助方法

張小剛* 高健斌* 高橋勇** 黒岩丈介** 小高知宏** 小倉久和**

* 福井大学工学研究科 ** 福井大学工学部

1 はじめに

現在、コンピュータへの入力装置としてキーボード、ライトペン、バーコードリーダなど様々あるが、やはりキーボードが最も有力な入力装置であると考えられる [1]。キーボードを用いた日本語入力方式には、主にカナ文字入力とローマ字入力がある。文字の読み方によるこの二種類の入力方式は、日本語の読み方が分かる人にとっては簡便であるが、日本語の読み方、特に日本語の漢字の読み方が分からない外国人日本語初心者にとっては大変難しい操作である。そこで、読み方が分からなくても漢字を効率的に入力できる方式の必要性がいわれている [2]。本報告では日本語の読み方によらない日本語漢字字画入力補助方法 JCSIM (Japanese Character Strokes Input Method) 法を提案し、提案システムの評価実験を行った結果を報告する。

2 JCSIM 入力方式

本報告では、日本語入力システムの補助方式、JCSIM 法を提案する。JCSIM 法では、全ての漢字字画を六種類に分ける。これらの六種類の字画と対応する六つのキーを用い、五回タイプして入力を終了する。画数が五画に足りない場合、実際の字画を入力してからスペースキーをタイプし、入力を終える。同じ入力コードには二つ以上の漢字が定義される場合では、全部の漢字に番号をつけて表示する。番号を押して、入力を終了する。

本方式では漢字の字画とキーを表 1 が示すように定義する。

表 1：漢字字画と入力キーの対応

漢字字画	一	丨	ノ	ㄥ	丿	その他
キー	S	D	F	J	K	L

Japanese character input auxiliary method by character strokes

Xiaogang Zhang* Jianbin Gao* Izamu Takahashi** Takesuke Kuroiwa** Tomohiro Odaka** Hisakazu Ogura**

* Graduate School of Engineering, Fukui University

** Faculty of Engineering, Fukui University

S(一)：左から右への水平的な字画；
D(丨)：上から下への垂直的な字画；
F(ノ)：右上から左下へまたは左下から右上への字画；

J(ㄥ)：左上から右下への字画；

K(丿)：字画“丨”の終点にはねがある字画 (はねの方向に無関係)；

L(その他)：上記以外のすべての字画。

表 2 と表 3 は「小倉研究室」を入力する際の入力字画例と JCSIM 法によるの入力キーである。

表 2：入力字画例

	第一画	第二画	第三画	第四画	最後画
小	J	ノ	ㄥ		
倉	ノ	ㄥ	一	その他	一
研	一	ノ	丨	その他	丨
究	ㄥ	ノ	その他	ノ	その他
室	ㄥ	ノ	その他	一	一

表 3：入力キー例

	第一画	第二画	第三画	第四画	最後画
小	K	F	J	スペース	
倉	F	J	S	L	S
研	S	F	D	L	D
究	J	F	L	F	L
室	J	F	L	S	S

3 システム構成

JCSIM 法にもとづく提案システムは入力、コード番号辞典検索、漢字検索と出力などの機能を持つ。その構成を図 1 に示す。

- 入力の受け取り：
ユーザからの入力を受け取って、一字としての入力が終了したかどうかを判断する。
- コード番号辞書検索：
入力字画による字画コード定義辞書からそれに対応する EUC コード番号を検索する。

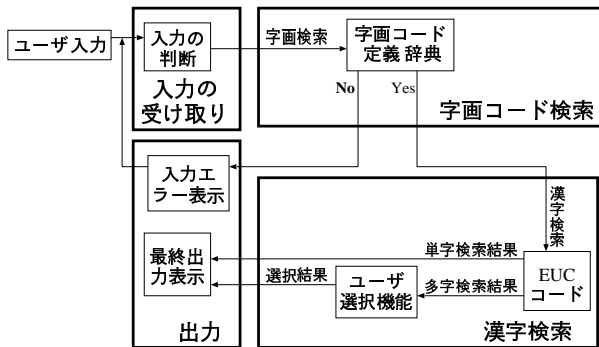


図 1: NKJ にもとづくシステム構成

● 漢字検索:

EUC コード番号による EUC コードから漢字を検索する。また単字検索の場合では検索結果を出力する。多字検索の場合、ユーザが選択してから出力する。

● 出力:

最終の漢字を出力する。入力エラーがあると、エラーメッセージを表示して新たに入力を待つ状態に戻る。

同じ入力コードに二つ以上の漢字が定義されている場合は各漢字に番号をつけて選択メッセージを表示する。

4 実験と評価

提案システムの有効性を確認するため、評価実験を行った。入力者は、外国人日本語初心者であり、実験を行う前に JCSIM 法の使用経験がない三人の留学生である。実験用データ (表 4 の「文」) は 100~200 文字以内の 3 種類の文とした。実験に関する詳細は表 4 に記す。

表 4 : 実験データ

文番号	文源	総字数	漢字数	仮名数
F 1	工業日本語	132	50	82
F 2	日本建築史	181	103	78
F 3	史籍雑纂	117	117	0

表 4 に基づく評価実験結果を図 2 に示す。図 2 の横軸は実験用文の漢字数を示すもので、縦軸はその文を入力する際の所用時間を表す。AR、BR、CR はそれぞれ入力者 A、B、C がローマ字入力方式を用いて 3 種類の文を入力に要する時間で、AN、

BN、CN は入力者は JCSIM 方式を使用した場合の所用時間である。JCSIM 入力法を用いることにより、ローマ字入力法に比べ、漢字を入力するのに要する時間を軽減可能であることが示された。

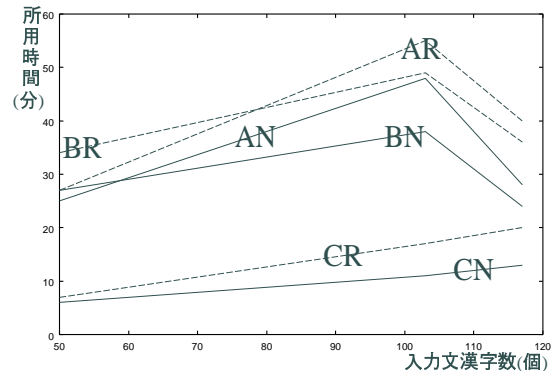


図 2: JCSIM 法とローマ字入力方式による実験結果

5 今後の課題

本手法を使用することにより、読めない漢字が効率的に入力可能である。しかし、漢字の字画は区別しにくく、間違っ分析される場合がある。また、入力者によって漢字の書き順も違う場合がある。これらの問題を解決するために、間違っ分析されやすい漢字については、正しい入力コードと間違いやすい入力コードも入力できるように用意することが考えられる。しかしながら、本手法では、同じ入力コードに二つ以上の漢字が定義されているものが多数存在する。これに加えて、間違っやすいコードまで加えると、入力コードに対応する漢字候補が多数存在し、正しい候補を一つにしぼり込む手続きを改良する必要が生じる。また、漢字の字画分析という負担をユーザにかける問題がある。これらの問題を解決するのが今後の課題である。

参考文献

[1] 尾関雅則, “第 32 回全国大会にあたって一日本語入力方式について”, 情報処理, Vol.27, No.5, pp493-497 (1986)

[2] 白井英俊, 荒木卓也, “日本語入力手段としての漢字モード”, 計量国語学, Vol.14, No.6, pp.239-251 (1984)