

レストランのメニュー画像からの文字領域切り出しと認識に関する研究

Study on Text Detection and Character Recognition from Menu Images

研究学生 中西 航 指導教員 江崎 修央

1. はじめに

近年、情景画像からの文字領域抽出・認識を利用したシステム開発が注目を集めており、旅行先での自動翻訳システムや、図書の自動データベース化などに利用されつつある。当研究室でも、これらのシステム開発に着目し、4年前から視覚障害者向けの情景画像中の文字情報読み上げシステムについて研究を行ってきた[1]。

現在、我々は実用化の第一歩として、レストランのメニュー読み上げシステムの開発を行っている[2]。本稿では、日本語メニュー画像からの文字領域抽出率の向上実験、投票制を基本とする文字認識精度の向上実験を行った。

2. メニュー画像からの文字領域抽出

2.1. 対象画像

今回対象としたメニューはインターネット上で取得した居酒屋のメニュー画像である。店舗数は26、画像数は79枚である。図1にメニュー画像の一例を示す。画像には文字情報だけではなく、メニュー画像なども掲載されている。



図1 メニュー画像の例 [3]

2.2. 文字領域の抽出

対象メニュー画像の文字サイズを調査した結果、今まで我々が対象としていた ICDAR2003 Robust Reading Competition[4]のデータに含まれる文字よりも、小さい文字がメニュー画像中に多く使われていることが判明した。そこで、これまでに提案してきた文字領域抽出手法の中で、小さな文字領域の抽出に有効である 8color2 手法と、B-Binary2 手法、Morphology 演算の応用手法である Closing-Opening を利用した手法を対象メニュー画像に適用することにした。

表1に結果を示す。文字領域の抽出能力の評価には、Precision=c/E, Recall=c/T (E: 候補文字領域, T: 正解文字領域, c: T と E との共通領域) を利用した。また、評価値 F は Precision と Recall の平均である。表中の “+” で表した結果は、複数の抽出手法で出力された領域の OR を取った領域であり、これにより B-Binary2 と Closing-Opening を組み合わせる事で、高い抽出率が得られることを確認した。

表1 各手法による文字領域抽出率

	Precision	Recall	F
8color2 (8)	69.4	44.1	53.9
B-Binary2 (B)	78.5	42.1	54.8
Closing-Opening(CO)	66.4	48.3	55.9
8 + CO	61.5	62.4	62.0
B + CO	66.1	61.4	63.7
8 + B	68.2	46.0	55.0
8 + B + CO	61.1	63.1	62.1

3. 投票制による文字認識

3.1. 投票制による文字認識処理の概要

文字領域抽出に用いた画像を含む48枚のメニュー画像から、文字領域を1文字ずつ手動で切り出し、認識処理を行った。文字画像の切り出しの一例を図2に示す。ここで、切り出した画像に複数の画像処理手法を組み合わせた前処理を行うと、同じ文字画像にもかかわらず、別の認識結果を得る可能性がある。つまり、元画像にそのまま認識処理を行い正しい結果が得られなかったとしても、適切な前処理を行うことで正しく認識される場合がある。このことから、本実験では前処理の違いを利用した投票制による文字認識処理を行った[2]。今回利用した OCR 処理系は AI ソフトの「活字文書 OCR ライブラリ」である[5]。このソフトウェアは、1つの画像に対して1つの候補文字列を出力するため、元画像を行う画像処理手法の組み合わせ数が総投票数となる。最終的に、投票結果の多い順から上位10候補を出力した。



図2 文字画像切り出し例

3.2. 認識結果

投票制による文字認識処理の結果が図3である。グラフの横軸は順位、縦軸は累積の認識率を表している。図中の菱形のマーカで表されるグラフは、投票制による文字認識を行った結果である。このグラフから1位ではまだ認識精度が低いものの、2位まで見ると認識精度は37%となり、10位までの累積結果で認識率は51%程度に上がることを確認した。次に四角のマーカは、小さい文字画像の高さを15ptに拡大して認識を行った結果。三角のマーカは、拡大処理に加え、大きすぎる文字画像の高さを15ptに縮小して認識を行った結果を表している。小さい文字画像を拡大することで、認識能力が10%向上することも確認した。

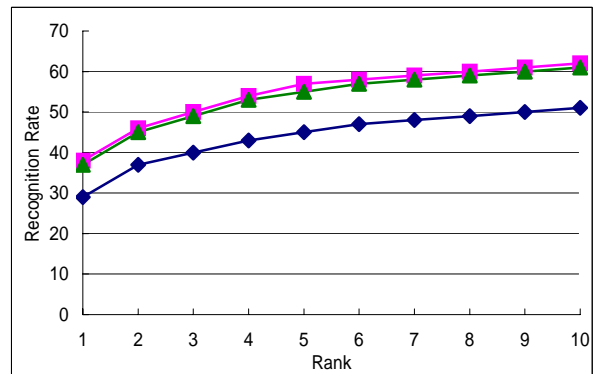


図3 投票制による文字認識処理の結果

参考文献

- [1] Nobuo Ezaki, Kimiyasu Kiyota, et al: Improved Text-Detection Methods for a Camera-based Text Reading System for Blind Persons, International conference on Document Anlysis and Recognition, (2005)
- [2] 江崎 修央, 砂崎 由樹, 中西 航, 清田 公保: 視覚障害者向けレストランメニュー画像からのメニュー読み上げシステム, 2006年2月23日メディア工学研究会にて発表予定
- [3] 「日本料理曾根崎乃多幸梅ホームページ」
<http://www.takoume-kankou.co.jp/>
- [4] S.M. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong, and R. Young, "ICDAR 2003 Robust Reading Competitions", Proc.of the ICDAR, pp. 682-687, (2003)
- [5] 「エー・アイ・ソフト」
<http://www.aisoft.co.jp/>