

ナンバープレート情報認識装置

2G-8

(漢字情報の認識)

森 和宏 内山 ありさ

(株)東芝 情報処理・機器技術研究所

1. はじめに

近年、駐車場の無人管理や旅行時間計測などの応用を目的として、「情景画像中の文字認識技術」によるナンバープレート認識が研究されている。

我々は先にナンバープレートの数字情報を実用的な認識率、処理時間で読み取る装置について報告した¹⁾が、応用によっては、漢字情報(陸運支局)も認識する必要がある。

漢字情報は高価な高解像度カメラを用いなければ一般的に解像度が低く、また照明条件の変動などから、画質が悪くなることもある。このような、低解像度、低品質の漢字情報を精度良く認識するために、以下のような手法を用いた。

- ・ ナンバープレート規約を知識とした、漢字領域の推定
- ・ 垂直、水平成分個々の連結領域解析による、接触の少ない検切り
- ・ 観測パターンを濃度圧縮して直接辞書と照合する観測パターン認識
- ・ 自動発生した、変形パターンの辞書への組み込み
- ・ 陸運支局名を知識とする後処理

2. 検出処理

(1)漢字領域の推定

ナンバープレートは規約²⁾によれば、図1の様な寸法で記載される。すなわち一連車両番号、ならびに分類番号の検出情報から、相対的に大まかな領域を推定することが出来る。

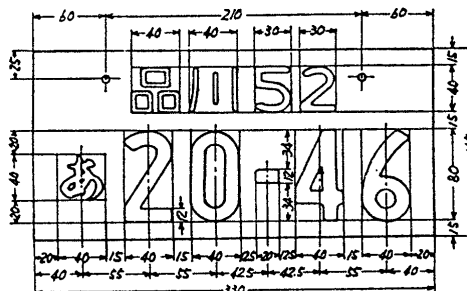


図1. ナンバープレート規約(中板)

(2)連結領域解析

推定された領域内の濃度分布を計測し、コントラストを改善する。

2値化は、一連車両番号と同様に、空間フィルタリングによる微分処理を施し、所定の閾値で2値化する。その際に、垂直微分、水平微分を独立に施し、図2に示すように文字の水平、垂直成分を個々に検出する。

(水平成分)

(垂直成分)



図2. 水平・垂直成分検出(品川の例)

個々に検出された文字の水平、垂直成分を各々ラベリング処理し、連結領域情報を解析して、文字らしい大きさとして組み合わせ可能な連結領域の集まりを文字候補として選出する。

A Vehicle License Plate Recognition System

Kazuhiro Mori, Arisa Uchiyama

Information Systems Engineering Lab. TOSHIBA Corp.

選出された全ての文字候補同士の位置関係を解析し、文字列として組み合わせ可能なものを文字列候補として選出する。

1～4文字の文字、文字列において各々定められた評価基準により代表を選出する。選出された代表は全て認識処理の対象となる。

3. 認識処理

(1) 観測パターン認識

ナンバープレートの漢字情報のように低解像度、低品質の文字を認識しようとする、2値化した時点での潰れや擦れがひどく、読み取り精度が著しく低下する。

そこで、文字パターンを2値化しないで、観測パターンをそのまま類似度演算に用いる観測パターン認識を用いた。図3に示すように、文字領域内で濃度分布を計測し、8bitから4bitに圧縮して辞書パターンとの積和演算を行う。これにより、低解像度、低品質の漢字パターンを精度良く認識できた。

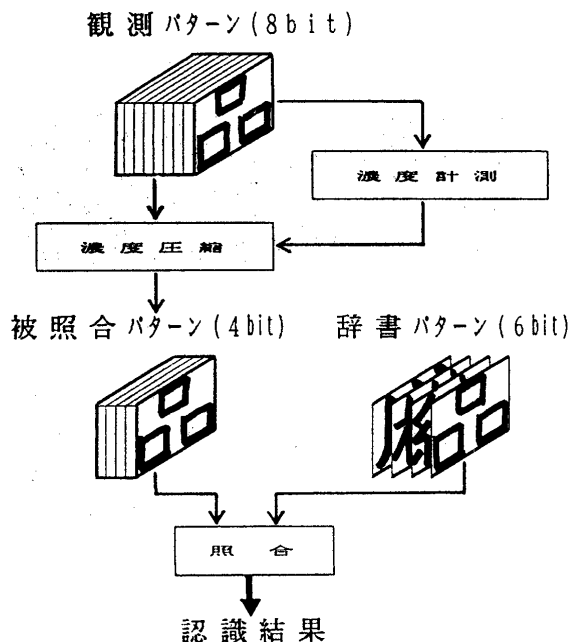


図3. 認識処理の流れ

(2) 辞書設計

観測パターンをそのまま辞書と照合する場合、大きさの正規化を行うことが難しい。画像中のプレートの大きさは極端には変わらないものの、多少の変動は生じる。したがって漢字パターンの大きさも多少変動する。またプレートの傾きなどによる変形も生じる。そこで辞書パターンを作成する際、大きさや傾きの変動による変形パターンを自動発生させ学習パターンに組み込むことで、大きさの変動や傾きによる変形にも対処している。

4. 評価結果

表1に1,414枚のナンバープレート画像に対する漢字情報の認識結果を示す。この場合の母数(データ数=1,414)は、一連車両番号と分類番号の両方が正しく検出できたデータの数である。また正解数は陸運支局文字列として正しく認識できた数である。

データ数	1,414
正解	1,348 (95.3%)

表1. 評価結果

5. おわりに

検切り処理の工夫、観測パターン認識、変形パターンの自動発生により、低解像度、低品質のナンバープレート漢字情報を精度良く読み取ることが出来た。今後さらに様々な実用環境下で評価、改良を行っていく。

参考文献

- 1) 森、他4名：「ナンバープレート情報認識装置(数字情報の認識)」, 第39回情処全大、pp.526-527、1989
- 2) 注解自動車六法 (株)第一法規出版