

検出容易な文字パターン生成に関する検討

山下 健司* 内田 誠一** 岩村 雅一*** 大町 真一郎**** 黄瀬 浩一***
(九州大学大学院*システム情報科学府 **システム情報科学研究所)
(***大阪府立大学大学院工学研究科) (****東北大学大学院工学研究科)

1 はじめに

情景画像中の文字を認識するためには、まず文字領域の検出が必要となる。従来、この検出処理には、文字パターンが一般的に持つ普遍的な特徴を用いられてきた。しかしどのような特徴を用いても例外的なパターンは存在し、高精度な検出は困難である。本研究では、検出容易な文字パターンを生成することで、この困難な文字検出問題の抜本的解決を目指す。

本稿において検出容易な文字パターンとは、何らかの特徴空間の中で、文字分布の中にあり、かつ非文字分布から最も遠いパターンとして定義される [1]。具体的には、特徴空間に局所方向ヒストグラム特徴を用い、文字分布をフォント画像の集合、非文字分布を非文字画像の集合により規定した上で、部分空間法に基づく検出容易性の評価尺度により、検出容易な文字を生成する。

2 検出容易な文字パターンの定義

2.1 局所方向ヒストグラム特長による特徴空間

上述のようにパターンを表現する特徴として、本稿では局所方向ヒストグラム特徴を用いる。文字のエッジの局所方向の分布で文字を表現するというこの特徴は、色や明るさで変化しにくい性質を持つ。同一のフォントでも様々な色で用いられるため、色(輝度値)の変化に頑健な特徴により文字を表現するのは合理的と言える。

2.2 部分空間を用いた評価尺度

検出容易性の評価尺度の計算には部分空間を用いる。具体的には、まず文字および非文字の各クラスについて学習パターンを収集し、それらについて主成分分析を行い、得られた上位主成分を正規直交基底とした部分空間により、各クラスのパターンの分布を近似表現する。本稿では非文字分布を約8万種類の非文字画像から、文字分布を308種類のフォント画像集合から決定している。そのあとであるパターン p の検出容易性を評価する評価値 D を、 p と非文字部分空間とのなす角(正準角) θ_N と文字部分空間とのなす角 θ_C から閾値 Θ を用いて以下の式により定義する。

$$D = \begin{cases} 0 & (\theta_N < \Theta) \\ \frac{\theta_N}{\theta_N + \theta_C} & (\theta_N \geq \Theta) \end{cases} \quad (1)$$

この評価値により非文字分布から遠く文字分布に近いほど検出容易と定義する。

3 実験

上述した評価尺度により、308種類のフォント画像集合から最も検出容易とされたものをカテゴリごとに選出し、その検出容易性を調べる実験を行った。選出された文字パターンを図1に示す。

3.1 一般的なフォントとの検出容易性の比較

選出された文字と一般的によく用いられるフォントである TimesNewRoman とで検出比較実験を行った。その実験結果が図2である。左が選出された文字で右が TimesNewRoman である。この結果から、選出された文字パターンがより検出容易であることが確認できる。



図1: 選出された検出容易な文字パターン

図2: TimesNewRoman との検出比較実験結果

3.2 様々な条件下での検出実験

情景画像中で考えられる幾何変換や複雑背景中で、選出された文字の検出容易性がどの程度保たれるかを検証した。その結果を図3と図4に示す。この結果からある程度の幾何変換や複雑背景上であっても、検出容易性が保持されることを確認した。



図3: 回転させた文字での検出実験結果



図4: 異なる文字背景での比較実験結果

4 まとめ

本稿では、検出容易文字を、非文字分布から遠く文字分布内にある文字と定義し、そしてこの遠さを分布間に定めた距離尺度により評価することで選出された文字の検出容易性を示した。今後の課題として、より様々な状況での文字の検出実験や、評価基準に可読性やカテゴリ分離度を加える、既存フォントからの文字の選出ではなく新たな文字の生成を行う、などが考えられる。

参考文献

- [1] 服部亮史, 内田誠一, 岩村雅一, 大町真一郎, 黄瀬浩一, “検出容易な文字パターン生成に関する検討,” 信学技報, PRMU2008-209, 2009年2月.