

複合類似度法の辞書データが低品質文字認識に与える影響 Effect of Dictionary of Multiple Similarity Method on Degraded Character Recognition

熊谷 将也[†] 片山 英昭[†] 池野 成高[†] 池本 幸平[†] 中田 健一[†]
Masaya Kumagai Hideaki Katayama Shigetaka Ikeno Kohei Ikemoto Ken'ichi Nakata

1. はじめに

近年、文字認識の実用化にあたり、信頼性向上という観点から、汚れた文字や一部が欠けた文字などの低品質文字を精度良く認識する技術が強く望まれている。一般的な低品質文字認識では、文字の外乱除去や 2 値化処理などの前処理で対象に応じた独自の工夫が必要になる場合が多い。しかし、このような個別処理は汎用性を損ない、適用条件を誤ることで誤認識する可能性もある^[1]。したがって、文字認識手法単独において汎用的かつ信頼性の高い手法を提案することが重要である。そこで、手書き文字やファクシミリなどの低品質文字認識に一般的に用いられる複合類似度法^[2]を用いて文字認識を行う。

複合類似度法では、認識に用いる辞書データが最終的な認識率に大きく影響する。辞書データは、複数の学習データから作成され、その学習データの数や種類により組み合わせが数多く存在する。そこで本研究では、学習データの数や種類が認識率に与える影響をそれぞれ検討し、高い認識率が得られる最適な辞書データを作成した。

2. 複合類似度法

複合類似度とは、ある文字に関する特徴平面と入力文字のベクトルが成す角度がどれだけ 0°に近いかという度合いのことである。複合類似度法の利点は、複数の学習データにより作成した辞書を用意するため、単純類似度が単一辞書との比較であるのに比べて、雑音に対して安定であり、柔軟な認識が可能となることである。そのため、本研究では、外乱要因の多い屋外でのナンバープレートの文字認識にも適用可能であると考えた。

3. 辞書データの作成方法

辞書データ作成のフロー図を図 1 に示す。ある文字に関する画像データを 50×50 pixel の大きさに正規化し、複数の学習データとして用意する。これらの学習データから辞書データを作成することにより、ボケた文字や解像度の低い文字、またはフォントが多少違う文字にも柔軟な認識が可能になる。

次に、それぞれの学習データを 1 次元のベクトルに変換し、各要素から定正準化およびノルムの正規化を行ったものを、

$$\mathbf{x}_i = [x_1, x_2, \dots, x_N]^T \quad (1)$$

と表す。また、それらの学習データを並べた行列を、

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k] \quad (2)$$

で表す。さらに、行列 \mathbf{X} についての自己相関行列 \mathbf{Q} を求める^[3]。

$$\mathbf{Q} = \mathbf{X}\mathbf{X}^T \quad (3)$$

ここで得られる行列 \mathbf{Q} は実対称行列であることから、本研究ではヤコビ法により固有値・固有ベクトルを算出する。また、算出時には、計算時間の短縮と精度の保持のため、なるべく関数計算を避けるようにウィルキンソンの改良公

式を用いる。ここで、固有値を大きい順に並び替え、それに対応する固有ベクトルも並び替えて、

$$\{\varphi_1, \varphi_2, \dots, \varphi_R\} \quad (4)$$

とおく。

複合類似度法では、これらの固有値・固有ベクトルの集合を辞書データとして用いて認識を行う。

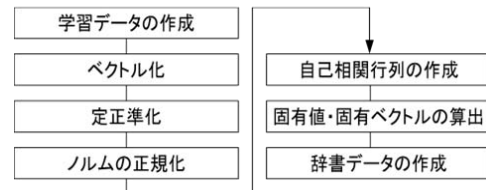


図 1 辞書データ作成のフロー図

4. 認識方法

入力画像の大きさを正規化し、一次元のベクトル化および定正準化を行ったベクトルを、

$$\mathbf{a} = [a_1, a_2, \dots, a_N]^T \quad (5)$$

と表す。また、(3)式の \mathbf{Q} の固有値を $\lambda_i (i = 1, 2, \dots, r)$ とすると、入力ベクトル \mathbf{a} の複合類似度は、

$$S(\mathbf{x}) = \frac{1}{\|\mathbf{a}\|^2} \sum_{i=1}^r \frac{\lambda_i}{\lambda_1} (\mathbf{a}, \varphi_i)^2 \quad (6)$$

で与えられる^[4]。ここで、 r は認識に用いる部分空間の次元数であり、特徴ベクトルの次元数を $n (= 2500)$ とすると $r < n$ である。

5. 実験方法

本研究では、光学ぼけや低解像度に焦点を当て、辞書データが認識結果に与える影響を検討する。評価は、誤認識が起こる可能性を調べるため、認識対象に対して最も一致度の高い文字と 2 番目に一致度の高い文字との一致度の差をとることにより行う。認識する対象は、「あ」から「ん」までの 46 文字のひらがなとする。ここで、46 文字の中で「る」と「ろ」の差が最も小さいため、本研究では「る」の辞書データを研究対象とする。また、辞書データの特性の検討は外乱の無い条件で行う必要があるため、辞書データに用いた学習データを認識対象としても使用した。

5.1 光学ぼけに関する辞書

本研究における光学ぼけ画像は、ガウシアンフィルタ処理を行った画像で仮定し、ぼけ度はガウシアンフィルタ処理を行う回数により決定する。本研究では、1 回から 12 回処理を行った 12 種類の画像を作成する。ここで、ガウシアンフィルタ処理を行った光学ぼけ画像の例を図 2 に示す。学習データの組合せと個数が認識結果に与える影響について検討する。

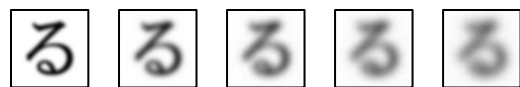


図 2 光学ぼけ学習データ

[†] 舞鶴工業高等専門学校, MNCT

5.2 低解像度に関する辞書

本研究における低解像度画像は、ニアレストネイバ法により解像度を低下させたものを使用する。ここで、ニアレストネイバ法により解像度を低下させた画像の例を図3に示す。学習データの個数が認識結果に与える影響について検討する。

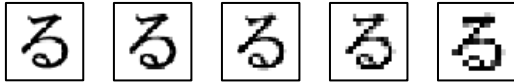


図3 低解像度学習データ

6. 実験結果

6.1 光学ぼけ学習データが認識結果に与える影響

ぼけ度の異なった12種類の画像から、連続した5枚の画像を選択し、作成した組合せにより認識させた結果を図4に示す。連続した5枚の画像の選択は、ぼけ度を1ずつ変化させたものとする。その結果、ぼけ度を高くするにしたがって、ぼけ度の低い対象に対する一致度の差は小さくなるが、ぼけ度の高い対象に対する柔軟性が確認できる。また、ぼけ度の連続した画像の個数を変化させ、認識させた結果を図5に示す。学習データとしてぼけ度の高い画像を多く取り入れることで、学習データの組合せに関する結果と同様、ぼけ度の低い対象における一致度の差は小さくなるが、ぼけ度の高い対象に対する柔軟性が確認できる。これらの結果から、光学ぼけに関する辞書データを、ガウシアンフィルタ処理を1回行った画像から5, 6回行った画像までの5, 6枚の学習データで作成することにより、光学ぼけに柔軟な認識が可能であると考えられる。

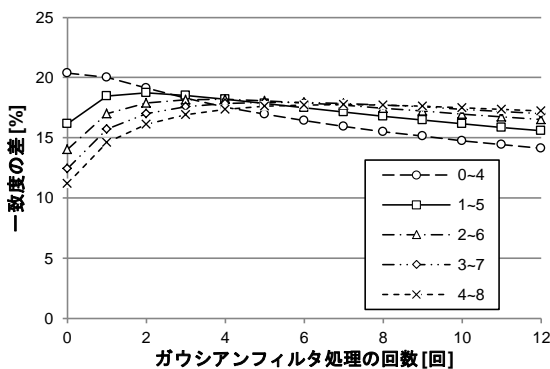


図4 光学ぼけ学習データの組合せ

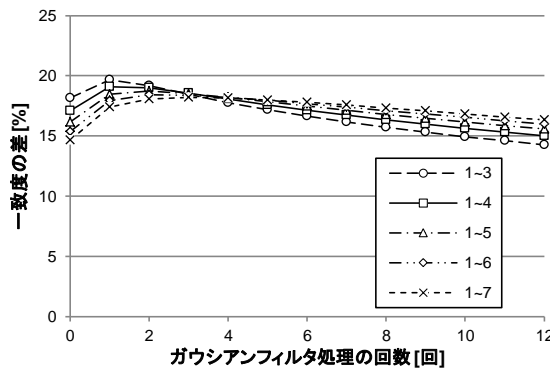


図5 光学ぼけ学習データの個数

6.2 低解像度学習データが認識結果に与える影響

低解像度学習データの個数を変化させ、認識させた結果を図6に示す。学習データとして解像度の低い画像を多く取り入れることで、 20×20 pixel以上の認識対象において柔軟性が確認できる。この結果から、低解像度に関する辞書データは、 20×20 pixelから 45×45 pixelの6枚の学習データで作成することにより、低解像度に柔軟な認識が可能であると考えられる(認識対象が 50×50 pixelの場合)。しかし、 20×20 pixelよりも低い解像度が認識対象である場合、どの辞書においても一致度の差が0付近となった。このことから、認識対象の解像度限界が 20×20 pixel程度であると考えられる。

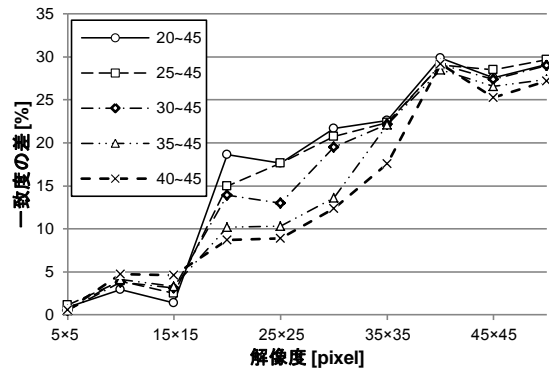


図6 低解像度学習データの個数

7. おわりに

光学ぼけ学習データが認識結果に与える影響を検討した結果、光学ぼけに関する辞書データは、ガウシアンフィルタ処理を1回行った画像から5, 6回行った画像までの5, 6枚の学習データで作成することにより、光学ぼけに柔軟な認識が可能であることがわかった。また、低解像度学習データが認識結果に与える影響を検討した結果、低解像度に関する辞書データは、 20×20 pixelから 45×45 pixelの6枚の学習データで作成することにより、低解像度に柔軟な認識が可能であることがわかった(認識対象が 50×50 pixelの場合)。さらに、認識対象の解像度限界は 20×20 pixel程度であることが確認できた。

今後の課題としては、ブレ等のようなその他の低品質文字に関する辞書データの検討、固有ベクトルの個数が認識結果に与える影響の検討などを行うことが挙げられる。これらを検討し、光学ぼけや低解像度、ブレなど全ての低品質文字に対して最適の一つの辞書を作成することで、低品質文字認識やナンバープレートの文字認識等に広く応用させることが可能である。

参考文献

- [1] 宮本 一正, 熊野 信太郎, 杉本 喜一, 玉川 光明, 英保 茂, “複数特徴量を用いた低品質定型文字の一認識手法”, 電子情報通信学会論文誌, Vol.J82-D-II No.4(1999), pp.771-779.
- [2] 谷口 慶治, 長谷博之, 画像処理工学応用事例編, 共立出版株式会社, (2005), p147.
- [3] 柳詰 進介, 目加田 慶人, 井手 一郎, 村瀬 洋, “携帯カメラによる動画像を用いた低解像度文字の認識手法”, 電子情報通信学会総合大会講演論文集(2004).
- [4] 大町 真一郎, 阿僧 弘具, “低品質文字認識におけるつぶれを補正する複合類似度法”, 信学技報, Vol.98 No.70(1998), pp.17-24.