

特徴点補間による曲線あてはめに基づく画像中の手書き字認識に関する検討

鈴木雅人†

北越大輔†

†東京工業高等専門学校情報工学科

1 はじめに

著者らは、画像中に含まれる手書き文字の認識を行う手段として、一般物体認識で用いられる局所特徴量を活用する方法を検討している。多様な字形を有する手書き文字を認識するため、局所特徴量を抽出する際に用いるスケール空間に上限を設け、文字を構成するストロークの端点を検出すると共に、そのストロークを直線または二次曲線で近似することにより文字の構造情報を抽出することが、認識精度改善に有効であることは実験により検証済みである [1]。しかし、ストロークの近似精度と認識精度との間には一定の相関があるため、例えば「近」「了」などの複雑な曲線を含む文字は、ストロークを十分な精度で近似することが難しく、従って、認識精度も低いのが現状である。本研究では、複雑な曲線の中に特徴点を補間することによって、従来は1本の曲線で近似していたストロークをより精度高く近似することにより、このような複雑な曲線を含む文字の認識精度を改善する手法について検討する。

2 認識アルゴリズム

日本語の手書き文字は、複数のストロークが複雑に交差し合っ一つの文字を形成するため、局所特徴量の一つである SIFT 特徴量 [2] を用いた認識アルゴリズムを構築する場合、Laplacian フィルタを適用する際のスケール σ は微小な値の範囲で変化させる必要がある。文字を構成するストロークの端点や屈折点では、 σ の変化に応じてその輝度はある特徴的な変化を示すため、本研究では、次の手順によってそのような特徴点を抽出する。

1. 対象画像に対し文字領域のラベリング処理を行う。
2. スケール σ の上限を $\sigma_{max} = 7.0$ とし、その範囲内で従来法を適用し特徴点を検出する。
3. 検出した各特徴点に対し、スケール σ に対する輝度の変化を検証し、端点や屈折点と推定される画素を特徴点として選定する。

A Study of Handwritten Chinese Character Recognition Method Based on Quadratic Curve Fitting Using Interpolation of Feature Points

†Masato SUZUKI †Daisuke KITAKOSHI

†Department of Computer Science, National Institute of Technology, Tokyo College

4. 冗長な特徴点の検出を防ぐため、特徴点どうしの距離が $d_{th} = 10(\text{pixel})$ 以下で同じラベルを持つ特徴点を削除し、その重心を改めて特徴点とする。

次に、同じラベルを持つ特徴点を、距離の近いものから順に選び出し、そのラベルを持つ黒画素と重なる度合いが高くなるように直線または二次曲線を配置する。このような直線または二次曲線のあてはめが可能であるとき、それを2つの特徴点の間のストロークの近似とし、文字の構造情報を抽出する。文字画像における特徴点どうしの距離は非常に小さいため、ストロークは原則として直線で近似するが、平均二乗誤差 E が 2.0 以上の場合には、二次曲線による近似を行う。曲線をあてはめるため、画像の輝度曲面を二次曲面で近似し、2つの特徴点を結ぶ稜線 [3] を抽出することにより投票点を算出し、ハフ変換を用いて曲線のあてはめを行う。

ここでは、二次曲線の1つとして放物線を考える。放物線

$$(y - c) = a(x - b)^2 \quad (1)$$

を1つ定めるためには3つのパラメータを定める必要があり、回転を含めるとパラメータ数は4となる。しかし、画素ごとに勾配情報 ξ を求めることができれば、

$$b = x - \frac{1}{2a}\xi \quad c = y - \frac{1}{4a}\xi^2 \quad (2)$$

により、パラメータ b, c は勾配 ξ と a を用いて計算可能である。また、この放物線を θ だけ回転したものを考えるとき、点 (x, y) における勾配方向の角度 ϕ は、

$$\xi = \tan\left(\phi - \frac{\pi}{2} - \theta\right) \quad (3)$$

でありから、ハフ空間は a, θ からなる二次元空間と考えることができる。

最後に、得られた特徴点に対して SIFT 特徴量を記述する。手書き文字のように字形が多様である場合、文字を構成するストロークの形や向き、ストロークどうしの位置関係には大きなばらつきがあるため、記述する特徴量に大きな違いが生じる可能性が高い。そのため本研究では、このような特徴量の違いによってマッチングが取れなかった特徴点に対し、識別時にスケールを動的に変化させて特徴量の再抽出を行い、多様な字形を吸収した特徴抽出および識別を行っている。

3 提案手法

文字を構成するストロークを二次曲線で近似するとき、例えば図1のような文字の場合には限界が生じる。あてはめる曲線の次数を上げるとそれに伴ってハフ空間

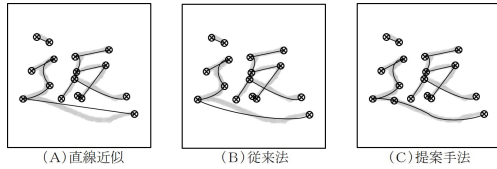


図 1: 二次曲線によるストローク近似

の次元も高くなるため、限られた特徴点による投票で曲線を特定するのは難しい。そこで本研究では、2つの特徴点の間に新たに特徴点を追加し、それぞれのストロークを直線または二次曲線で近似する方法を提案する。2つの特徴点 \vec{p}_0, \vec{p}_n の間の投票点列 $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_{n-1}$ に対し座標の平滑化処理を行い、各 \vec{p}_k に対して

$$\theta_k = \frac{(\vec{p}_k - \vec{p}_{k-1})^t (\vec{p}_{k+1} - \vec{p}_k)}{\|(\vec{p}_k - \vec{p}_{k-1})\| \cdot \|(\vec{p}_{k+1} - \vec{p}_k)\|} \quad (4)$$

$(k = 1, 2, \dots, n-1)$

を求め、 θ_k が最小となる \vec{p}_k を新たな特徴点として登録する。

4 検証実験

提案手法の効果を確認するため、手書き文字データを用いて実験を行った。対象データは著者らが独自に収集した 1148 字種 50 セット [4] で、字種の内訳は平仮名 71 字種・カタカナ 71 字種・教育漢字 1006 字種である。参照文字画像は一辺が 256(pixel) の正方形画像であり、認識対象画像は 4608(pixel) × 3456(pixel) サイズの風景画像である。あらかじめ互いに異なる 10 個の文字画像を 20 セット用意し、拡大縮小・回転の変換をランダムに作用させ、これらを風景画像に 10 個ずつ埋め込んだ。埋め込んだ文字は、平仮名 20 字種、片仮名 20 字種、漢字 160 字種である。

参照文字フォントとして、MS ゴシック体、MS 明朝体、手書き文字を用いた場合の認識結果を表 1 に示す。いずれの手法においても、明朝体を参照画像として用いた場合に認識精度が最も高く、提案手法によって認識精度を 3.5 ポイント改善することができた。提案手法により新たに認識できた文字は、「返」や「了」などの複雑な曲線構造を持つものが多いことから、複雑な曲線に対する特徴点の補間は、従来法の改善に有効であると言える。しかし、文字を構成するストロークそ

表 1: 認識実験結果

認識手法	MS ゴシック体	MS 明朝体	手書き文字
従来法	45.3%	62.6%	60.4%
提案法	43.6%	64.2%	63.9%

のものは非常に似通ったものが多いため、全体の認識精度は 6 割程度に留っている。更なる改善をはかるためには、それぞれのストロークの組み合わせに関する大局的情報の抽出が必要であると考えている。

5 まとめ

本稿では、複雑な曲線の間に特徴点を補間することにより、従来は 1 本の曲線で近似していたストロークをより精度良く近似し、そのような複雑な曲線を含む文字の認識精度改善について検討した。検証実験では、認識精度そのものは数パーセントの改善率に留まるが、提案手法により複雑な曲線構造を持つ文字の認識精度が大幅に改善された。今後は、文字を構成するストロークだけでなく、文字の大局的構造に着目した文字構造情報の抽出を試み、大幅な認識精度改善に繋がりたいと考えている。

尚、本研究の一部は科学研究費補助金 (基盤研究 (C) 課題番号 18K02971) の助成によるものである。

参考文献

- [1] 鈴木雅人, 北越大輔, “線分端点検出及び二次曲面あてはめに基づく画像中の手書き文字認識に関する検討,” 信学技報 PRMU2018-44, pp.55-60, Sept. 2018.
- [2] David G.Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal of Computer vision, vol.60-2, pp.91-110, 2004.
- [3] 長谷川周, 鈴木道孝, 塚本新, 伊藤彰義, “手書き文字認識の稜線特徴量における稜線強度形状の評価と検討,” 信学技報 PRMU2015-153, pp.115-120, Feb. 2016.
- [4] 鎌形周平, 鈴木雅人, 北越大輔, “低品質文字を用いた標準パターン辞書構築による手書き署名認識法,” 第 76 回情処全大, 1Q-4, Mar. 2014.