

手書き文字の特徴を活かしたフォント生成システムの提案

Proposal of font generating system maintaining the feature of handwriting

小林清哉^{*1}
Seiya Kobayashi

井上聡^{*1*2}
Satoru Inoue

^{*1} 埼玉工業大学大学院
Graduate School of Engineering, Saitama Institute of Technology

^{*2} 埼玉工業大学
Saitama Institute of Technology

For generating the Japanese fonts keeping the characteristic of handwriting, all fonts have to be handwritten. It means that enormous amount of time and effort is required. In this research, we propose the automatic font generating system by coupling the elements composing the characters such as "Flick Stroke, Pullout Stroke, etc." extracted from representative Japanese characters.

1. 序論

1.1 背景と目的

手書き文字の特徴を活かした独自の日本語フォントを作成するためには、すべてのフォントを一文字ずつ手書きで作成していかなければならず、膨大な時間や労力が必要となる。そのため現在までに様々な手書き文字フォント作成のための手法が提案されている。例として、書き手にいくつかの文字を構成する要素を記入してもらい、文字の特徴の抽出を行う。そこからフォントの生成を行う方法[中西 09][田丸 03]や、複数の文字から書き手の特徴を判別し、フォントの生成を行う方法などが挙げられる[安本 97]。しかしそういった手法を用いたフォント生成は同じ入力に対して出力される文字が常に一定のフォントになる。しかし、実際に人間が書く文字は一つとして同じ形状ではない。そこで本研究では、代表的な文字から「払い、とめ、跳ね」などの、文字を構成する要素を分解して抽出し、それらの作業の効率化と出力の度に微小に変化するフォントの生成を目的とした研究を行った。

1.2 筆画

文字を構成するにあたり、特別な情報を所持している漢字として「永」(永字八法)が挙げられる。しかしそれは「書」を書く際に必要な技法が八つ集まっているというだけで文字を構成する要素としては成りえない。そこで注目したのが筆画である。筆画とは文字を構成するための最小限の単位のこと、横、縦(縦)、点、堤(右上斜めはね)、撇(左斜めはらい)、捺(右斜めはらい)、鉤(横画または縦画からのね)、折の永字八法とは少し異なる八つの基礎とその変形文字を足したものであり、組み合わせ次第ですべての漢字が網羅できるものである。(図 1)

2. 研究概要

本研究では書き手の文字の筆画を抽出するフェーズとその特徴からフォントを生成するフェーズに分かれており、前者では書き手が記入した文字からフォント生成に必要な要素を取り出し、ベクトルフォントに対応できるような特徴抽出を行う。後者では事前に準備したテンプレートと比較し筆画の開始位置とベクトルを決定する処理を行っている。



図 1. 筆画

2.1 筆画の抽出

本研究ではフォントの生成を行うための文字を構成する要素となる八つの筆画を主とした特徴の抽出を行う。書き手には「永」の字を含む特定の文字を一文字ずつ記入してもらいフォント生成に必要な筆画の形状を抽出する。そのアルゴリズムとして、始めに、記入された文字の最短矩形を求め、最も小さい値を持つ位置座標上に存在する筆画から x 軸方向に 1 ピクセルずつ筆画の上部の探索を行い、10 ピクセルの範囲まで求めた所でその特徴点として付加する。また、筆画のベクトルを求めるために近似直線(線形近似)を出す、そこから垂線を下ろし、筆画の下部に到達したところで特徴点として追加する。以上の作業を筆画の終点まで行い、下ろした垂線の中心を近似曲線で繋ぎあわせることにより始点と終点の座標を取得する。しかしこのままでは筆画同士が重なっている場合、正確に形状の抽出ができない。そこで、上部と下部の特徴点のユークリッド距離が前回までの平均と比べ二倍以上の距離が離れていた場合と前回の近似直線と比べ傾きが 35 度以上の差が生まれた場合に関しては処理を行わず、前者は次と前の特徴点から算出したユークリッド距離を参考に仮の特徴点を付加する。(図 2)後者は近似直線上で 30 度以内に存在する特徴点の探索を行うことで筆画同士の重なりを無視できる。以上のことからドットから出力するフォントでは難しい伸縮に対応したベクトルフォントの生成が可能となる。図 2 では上記のイメージ図を添付する。赤い線は示す先に特徴点が存在し、青の線はユークリッド距離が二倍以上離れていたため取得しない線である。最後に緑の線の先には筆画の始点と終点の座標が示される。

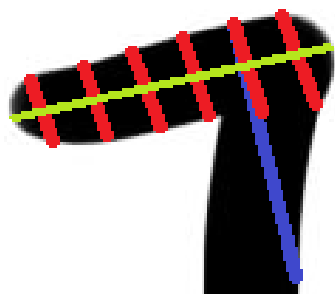


図 2. 筆画抽出のイメージ

2.2 テンプレートの作成

書き手の筆画から記入された文字とは違う、別の文字を出力するために、お手本となるテンプレートの作成を事前に行う。書き手の文字の特徴に最も近似しているフォントの生成を行うために、テンプレートを多量に用意することで、より、書き手の文字の特徴に近似したフォントの生成が期待できる。テンプレートの作成方法は、各文字に対して作成者が五回の記入を行い、筆画の開始と終わりの位置座標と、それらが示す二次元ガウス関数の確率密度(1)の作成、そして、筆画の長さの四つの要素を保持することでテンプレートの完成となる。

2.3 確率密度関数(PDF)

確率密度関数(Probability Density Function)とは連続の確率変数が取り得る値であり、一つ一つの値に意味はなく、範囲 $[a, b]$ 内の値をとる確率 $\Pr(a < X < b)$ を指定して積分することにより初めて意味を持つ。また、全範囲の確率は合計して 1(100%)となるように指定する必要がある。本研究で用いる二次元座標上の確率密度の場合は(1)式で求められる。また、 μ は平均値で σ は標準偏差である。

$$f(x, y) = \frac{1}{2\pi\sqrt{\sigma^2}} \exp\left(-\frac{\left(\frac{x-\mu_1}{\sigma}\right)^2 + \left(\frac{y-\mu_2}{\sigma}\right)^2}{2}\right) \dots (1)$$

3. フォント生成

特徴抽出フェーズで書き手が記入した後、各筆画の始まりと終わり、そして長さの三つ要素の平均と比較し最も近似しているテンプレートが選択される。そして、そのテンプレートが持つ座標の位置情報を参考に作成した二次元ガウス関数の確率密度から筆画のスタート位置が決定される。また、どの方向に筆画を伸ばすのか決定づけるため、テンプレートの終点座標を参照しベクトルを指定する。その後、書き手の筆画の形状を当てはめることによりフォントの生成を行う。

4. まとめ

本研究では出力の度に微小に変化し、且つ書き手の負担を軽減したフォント生成を目的とした。初めに、特徴抽出フェーズでは近似直線を用いて求めていくのだが、一部の文字に対して正確に取得できない問題が発生した。そのため、実験段階では書き手は文字を直接一画ずつ記入し、各筆画の形状を保存することにより特徴の抽出とした。次にフォント生成時では、筆画の始点として選ばれる座標に、テンプレートが保持している x 軸と

y 軸座標上の最も端に出力した特徴点がほとんど出力されない問題が生じたため、二次元ガウス関数の確率分布の範囲 $[a, b]$ 内の値をとる確率を、各軸上の最も小さな値 a 、最も大きな値 b と定義したとき、各値に $(b-a)/2$ を足すことにより図 3 のガウス分布のように、端の点の座標が出力される可能性を増幅させた。そのことにより、目的である出力の度に微小に変化するフォントの生成を行うことが可能となる。

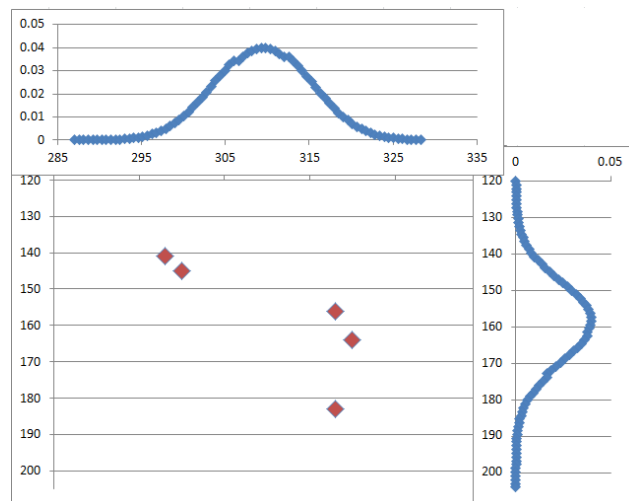


図 3. 二次元ガウス関数の確率密度分布

5. 今後の展望

現状として、一部の筆画の抽出と書き手の筆画の形状を繋げてフォントを生成するに至っておらず、簡易的な筆画抽出を行っている。また、テンプレートとして使用するデータ量が不足しており、満足のいく結果が得られていないことが挙げられる。今後の展望としては上記で挙げた問題の改善と、新たなアルゴリズムとして、書き手の年齢、性別が異なる場合に文字の特徴も異なると考え、テンプレートと比較するアルゴリズムの変更と、そのような情報からもフォントの形状を微小に変化させることを考慮したフォントの生成システムの提案とサンプルの収集を行おうと考えている。

参考文献

- [中西 09] 中西貴之, 荒木修: 書き癖パラメータ設定による手書き風フォントの自動生成, 映像情報メディア学会技術報告, 2009.
- [安本 97] 安本護, 池田尚志, 堀井洋: 大域的個性と局所的個性に基づく手書き風フォントの生成, 電子情報通信学会論文誌, 1997.
- [森 08] 森哲哉, 飯塚重善, 杉山精, 徳永幸生: イメージマップを用いた感性を反映するフォント生成システム, 全国大会講演論文集, 2008.
- [堀田 07] 堀田創, 野澤貴, 萩原将文: 感性ルールベースを用いた日本語フォント自動作成システム, 情報処理学会論文誌, 2007.
- [田丸 03] 田丸雅純, 村上真, 杉本富利, 米山正秀: パーソナルフォント生成のための文字画像の特徴抽出, 電子情報通信学会ソサイエティ大会講演論文集, 2003.