

統計的特徴量変換に基づく日本語手書き漢字変換

大垣慶介¹齋藤大輔²峯松信明³広瀬啓吉³東京大学工学部電子情報工学科¹東京大学大学院工学系研究科²東京大学大学院情報理工学系研究科³

1 はじめに

我々が文章に接する際、文字の字体によってその印象が大きく変わることがある。近年では計算機による印刷物や電子文書に接する機会が増え、その役割を計算機の文字インタフェースであるデジタルフォントが担うようになってきた。最近では、高速道路の標識にデジタルフォントが採用されるなど、その重要性は増してきている。フォントに求められる要素として、可読性の他、表現力が挙げられる。アニメーションや本の装丁などではこの側面が重要視されており、ウェブなどにおいても今後表現力を重視したフォントの需要は増加すると考えられる。このような表現力を重視したフォントとして、手書きフォントが用いられることがある。

しかし日本語漢字には 6,355 文字 (JIS X0208) の文字種があり、手書きフォント作成のためにこのすべての漢字を筆記することは多大な労力を要する。ゆえに個人のウェブページや手紙などで、ユーザー本人の個性を反映した手書きフォントを用いたい場合に、それを作成するのは困難である。そこで、文字に現れるユーザーの個性を表現するモデルを少数のサンプルから学習し、それを他の文字に適用するなどの手法が必要となる。しかし現在このような日本語手書き文字変換の研究は少ない。

ストローク分離が容易なオンライン手書き入力からの文字形変換を考えると、これは時系列で現れる点列のデータであり、同じく時系列を扱う音声工学と関連深い。音声の研究では近年隠れマルコフモデル (Hidden Markov Model; HMM) のような確率統計モデルを用いた手法が盛んに研究されている [1]。文字認識の分野でも音声認識の手法を参考に HMM を用いた認識手法の研究が行われている [2]。手書き漢字変換は、入力として標準フォントなどから字形をあたえ、それに個性を付与する形で手書き漢字へと変換する。これは、入力音声に目的話者の話者性を付与する声質変換と対応付けられる。

本研究では声質変換に広く用いられる混合ガウスモデル (Gaussian Mixture Model; GMM) に基づく統計的特徴量変換を利用し、入力文字の各ストロークに目的筆者の個性を付与する手法を検討する。

2 GMM に基づく統計的特徴量変換

統計的特徴量変換では入力、出力の D 次元特徴量ベクトルをそれぞれ x, y とし、これら 2 つの連結特徴量ベクトル $z = [x^T, y^T]^T$ の結合確率密度分布を推定する。推定した分布と、新たな入力 x から、出力の分布 $P(y|x)$ を求める。この入力からの出力 \hat{y} は期待値として求まる。結合確率密度分布として GMM を用いると、分布のパラメータを $\lambda^{(z)}$ とし、出力の分布が

$$P(y_t|x_t, \lambda^{(z)}) = \sum_{m=1}^M P(m|x_t, \lambda^{(z)})P(y_t|x_t, m, \lambda^{(z)})$$

と表される。以降この GMM を用いた特徴量変換の、手書き文字合成への応用について述べる。

3 日本語手書き漢字変換の提案手法

個人性を付与する変換を考える上で、特徴量の選択が重要となる。各文字を画像としてとらえ、全体に線形変換を施すと、傾き・大きさの個人性を付与することができる。これは、文字の重心を原点とした位置ベクトルを特徴量として用い、分布を単一正規分布で推定した場合の統計的特徴量変換によって表される。

しかし漢字は、偏旁冠脚[†]、ストロークに分解され、各ストロークは点列として表現される (階層的構成)。このことから、先の HMM 文字認識や、漢字と同様に部分要素の組み合わせから一文字がなるハングルの文字合成 [3] では文字を階層化し、階層ごとの特徴を用いるアプローチをとっている。本研究でも手書き漢字の合成問題を階層化し、正規化ストローク中の点列の座標変換によるストローク形状決定と、ストロークの配置と大きさの変換による文字生成の 2 段階に問題を分割した。

3.1 ストローク形状の変換

この節では始点を $(0, 0)$ に固定し、経路長を 1 に正規化したストロークの形状変換を述べる。このストローク中の各点の座標を変換することで目的筆者の正規化ストローク形状を得る。変換には座標 (x, y) に加えて、近傍の点との差分である動的特徴量 $(\Delta x, \Delta y)$ を用いる。動的特徴量は各点のストロークの局所的な方向を表し、変換後のストロークを平滑化する効果が期待される。声質変換の分野で動的特徴量を用いて品質を向上させた研究がある [4]。変換のための対応付けは、入出力ストロークの経路長を等分割した点とその順で対応しているとするもの、及び DP マッチングを用いるものの 2 種類について検討

[†]部首などの、漢字の左右上下に分解できる構成部分

Japanese Kanji handwriting conversion based on statistical feature transformation

¹Keisuke Ogaki, ²Daisuke Saito, ³Nobuaki Minematsu and ³Keikichi Hirose,

¹Faculty of Engineering, the Univ. of Tokyo, ²Graduate School of Engineering, The Univ. of Tokyo ³Graduate School of Information Science and Technology, The Univ. of Tokyo

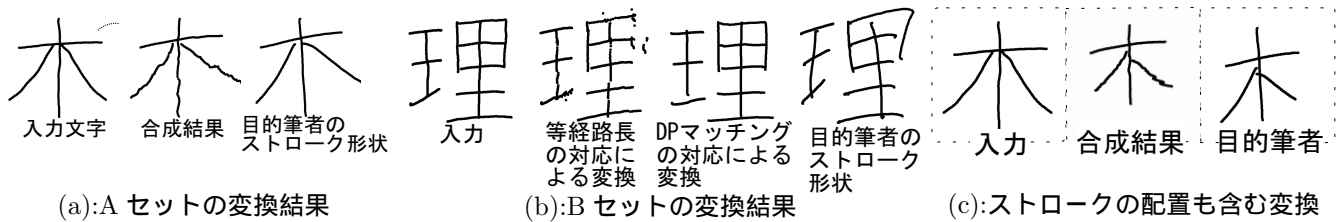


図 1: 実験結果

する。等分割に基づく対応では、点列の間隔が入出力で一定となり、合成時に Δ と座標の情報を容易に複合的に用いることができる。しかし一方で、折れの前後の点が対応するなど、特徴量の大きく異なる点が対応付けられる可能性がある。DP マッチングは、入力と目的筆者の、座標と Δ それぞれの差分をペナルティとしてその経路に沿った和がもっとも小さくなる経路を求め、ミスマッチの少ない対応付けを作る。しかし DP マッチングを用いると、等分割の対応付けと異なり、出力点列は等間隔にならない。そこで点列の間隔によらない出力特徴量として Δ の角度 θ を用いる。 θ と正規化された入力点列の間隔により正規化された変換ストロークが求まる。

3.2 ストロークの配置、大きさの決定

この節では、入力ストローク $x = [x_0, \dots, x_N]$ から目的筆者ストローク $y = [y_0, \dots, y_N]$ の始点 y_0 と大きさ $|y|$ を求め、前節の正規化目的筆者ストローク (y' とおく) を適切に配置する手法を述べる。ストロークの大きさ、配置の癖がともに終点 x_N の方向に影響を受けると仮定し、その正規化座標を入力特徴量とする。始点の変換は両筆者の始点の差 s を出力特徴量に用いる。大きさの変換は、目的筆者のストローク終点 y_N を入力ストロークの大きさ $|x|$ を用いて正規化した $\frac{y_N - y_0}{|x|}$ を出力特徴量に用いる。この変換による推定値から、両筆者のストロークの大きさの比 R が求まる。求めた s 、 R 、正規化目的筆者ストローク y' を用いて出力ストローク \hat{y} の生成式は以下で表される。

$$\hat{y} = Ay' + b \quad \text{ただし } A = R|x|, b = x_0 + s$$

4 実験

提案する、ストローク形状の変換、および配置と大きさの変換法の効果を調べた。変換元を楷書フォントの文字とし、目的筆者1名の手書きへの変換実験を行った。ただし楷書フォントの点列として、タブレット PC を用いてフォントの実際の文字の中心をなぞって得た点列を用いる。実験には二種類の文字セットを用いた。1つは、文字を構成するストロークがほぼ直線のみ漢字 29 文字の A セット[†]、もう1つは平仮名・片仮名・複雑な漢字を含む 28 文字の B セット[‡]である。

ストローク形状の変換に関して、1) 等分割による対応付け、2) DP による対応付けの効果を検証した。等分割の

対応づけによる変換の有用性を確かめる実験を行った結果が図 1(a) である。A セットのうち 28 文字で学習して、29 文字目の“木”に変換を行った。形状の比較のため、比較対象には目的筆者の各ストロークの始点と大きさを入力に合わせてシフトしたものをを用いている。合成結果では各ストロークの癖が目的筆者に近づいている。

DP による対応付けの効果を確認するため、より複雑な B セットを用い、等分割・DP の両対応付けの比較を行った結果が図 1(b) である。27 文字で学習して、28 文字目の“理”に変換を行った。先と同様、シフトさせた目的筆者文字を比較対象としている。DP マッチングにより対応付けのずれが解消され、字形の崩れが見られなくなった。しかし一方で各ストロークが楷書の直線的なものからあまり変化していない結果となった。この結果から、DP マッチングを用いた上で動的特徴量と正規化座標を複合的に用いるなど、より個人性を表せる特徴量や変換法への検討は今後の課題となる。

最後に、正規化ストロークに始点、大きさの変換を施す実験を A セットについて行った。先と同様 28 文字で学習して、“木”の変換を行った。ストローク形状は等経路長の対応により変換している。結果は図 1(c) である。これを見ると、大きさと配置がともに筆者のデータに近づいたといえる。

5 まとめ

本論文では手書き文字合成の手法として、統計的特徴量変換を用いて楷書のストロークから手書きストロークに変換する手法を提案した。この手法で簡単な漢字に関して手書き風の文字を生成できることを示した。今後、公開されている手書き漢字データベースで文字の変換を行い、評価実験を行いたい。また、複雑な漢字の個人性の表現は課題として残る。

参考文献

- [1] K.Tokuda *et al.*, Proceedings of 2002 IEEE Workshop on Speech Synthesis, pp.227–230 (2002)
- [2] 嵯峨山茂樹他, 信学技報, PRMU2000-35 (2000)
- [3] H. Choi *et al.*, IEICE Trans. on Information and Systems, vol.92, no.4, pp.653–661 (2009)
- [4] T.Toda *et al.*, IEEE Trans. on Audio, Speech, and Language Processing, vol.15, no.8, pp.2222–2235 (2008)

[†]井、王、交、左、作、仕、止、十、上など
[‡]ぬ、ほ、ユ、愛、雅、響、激、結、誤など