

文字認識器の学習データ自動生成ルールの獲得 Synthetic Handwriting Generation Rules for Handwriting Recognition Systems

福留 拓也
FUKUDOME, Takuya

概要: 本論文では、手書き文字の認識器の精度向上を目的とした学習データの自動生成手法について述べる。手書き文字認識器の学習に効果的なサンプルデータの自動生成手法が提案されているが、学習のノイズとなる不自然なデータが生成される場合もあり、認識器の精度向上が頭打ちになっている。本研究では、自然に見える文字の生成を目的とし、生物の形態進化シミュレーションのアイデアを用いた **Biohandwriting** による手書き文字の自動生成手法を提案する。生成された文字画像を評価した結果、文字認識器の学習データとして利用するには多様性が不十分であったが、従来手法よりも不自然な文字の生成を約 20%抑えることができた。

Summary: This article first describes synthetic handwriting generation for handwriting recognition systems. There are reported some methods for synthetic handwriting generation, and they can show good performance to generate synthetic handwriting as additional training dataset. But, some synthetic data are seemed unnatural. These unnatural data will be noise for training recognition systems. Therefore the improvement of the synthetic generation performance to enlarge training dataset for handwriting recognition systems, reaches the ceiling. In this research, I propose a method of synthetic handwriting generation with “Biohandwriting” for generating synthetic data seemed natural. It generates dataset by selection natural data, and unnatural data will be reduced. Finally, I show results of the experiment that compare generated synthetic data by the proposed method with randomly synthetic data.

キーワード: 手書き文字認識・文字生成・遺伝的アルゴリズム

Keywords: Handwriting Recognition, Synthetic Handwriting Generation, Genetic Algorithm

1. はじめに

文字認識は、コンピュータサイエンスにおいて古くから研究されているテーマの一つである。現在では、パターン認識技術の発展により郵便番号の自動読み取りや、活字文書の読み取りなど対象を限定した文字認識の商業利用が進んでいる。しかし、商業システムは、厳しい制約や専門的な知識が必要であり、人間と同等の認識能力を得るに至っていない。一方で、制限のない手書き文字の認識は、オフィスオートメーションやデジタルライブラリのようなアプリケーションへの応用が考えられる挑戦的な課題である。一般に認識器の性能は、認識のアルゴリズム・特徴抽出の方法・学習サンプルの量に依存する。特に、学習サンプルの量が認識の精度に与える影響は大きく、実用的な認識器を構築するためにはサンプルとなるデータの収集が重要である。しかし、サンプルデータの収集はコストがかかり、学習に十分なデータの収集は困難である。

手書き文字の認識器の学習において、データの不足を補うために、手書き文字を自動生成してデータセットを拡張する手法がいくつか提案されている。自動生成した文字データをデータセットに加えて学習することで、文

字認識器の性能向上が示された。しかし、生成された文字画像には不自然な文字が含まれる場合もあり、認識精度の向上は頭打ちになっていると考えられる。本研究では、手書き文字認識器の学習データの自動生成を目的として、生物の形態進化シミュレーションのアイデアを用いた **Biohandwriting** を提案し、自然な手書き文字の自動生成手法を目指す。提案手法により生成した文字画像を、形状の自然さとばらつきの観点から評価する。また、似た文字同士の特徴量を比較することで、提案手法により自動生成した文字が類似文字に共通の特徴をもつことを示す。

2. 手書き文字の自動生成手法

手書き文字の生成手法として、手書き文字の画像の幾何変換によって新しい文字画像を生成する手法と、文字のプロトタイプと筆記動作のモデルにより文字画像を生成する手法が提案されている。幾何変換により新しい文字画像を生成する手法は、自然な文字画像を生成することができるが、生成される文字画像は変換元に依存するため多様性が低いと考えられる。また、新しい文字画像を生成するために変換元となる手書き文字画像が必要である。一方で、文字のプロトタイプと筆記動作のモデル

この研究の一部は、電子情報通信学会の学生会研究発表会において発表の予定である。

によって、新しい文字画像を生成する手法では、文字のプロトタイプをモデルに基づいて変動させることで様々な文字画像を生成することが可能である。また、文字のプロトタイプを用意できれば、人手による手書き文字の画像が必要なく、幾何変換による手法と比較して文字生成のコストが低いと考えられる。

Tamás らは、ベジエ曲線による文字のテンプレートと腕の運動モデルである Delta-Log Normal Model による文字の自動生成手法を提案し、生成した文字画像が手書き文字認識器の学習に効果的であることを示している[1]。Delta-Log Normal Model は人間の腕の動作速度に関する筋神経の働きをモデル化した理論である。単一の弧を描く運動の速度は、距離や曲率、傾き、神経の応答時間・反応時間などの9つのパラメータにより再現される。文字の筆記のような複雑な軌道を描く運動も、複数のストロークの重ね合わせによって再現されることが示されている[2]。Tamás らの手法では、ベジエ曲線によるテンプレートから仮想的にストロークを抽出し、Delta-Log Normal Model による文字の軌道を再現している。テンプレートとなるベジエ曲線の制御点パラメータや、Delta-Log Normal Model の各パラメータをランダムに変動させることで、多様な形状の文字が自動生成される。しかし、人手による手書き文字にみられない不自然な文字を含む場合もある。自然な文字と不自然な文字の例として、文字“a”と“b”について、自然と判断される文字を図1(a)に、不自然と判断される文字を図1(b)に示す。文字認識器の学習に効果的な文字画像を生成するためには、不自然な変動を抑制する必要がある。

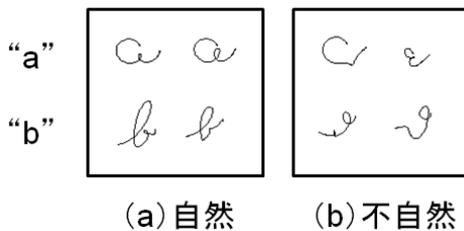


図1 自動生成された文字の例

3. 提案手法

文字認識器の学習に効果的な文字画像の生成を目指して、生物の形態進化のシミュレーションのアイデアを用いた Biohandwriting を提案する。手書き文字の自動生成手法として、文字のテンプレートと筆記運動のモデルによる手法を基に、Biohandwriting による自然な文字画像の生成を目指す。

3.1. 形態進化のシミュレーションによる文字生成

自然な手書き文字の生成手法として、生物の形態進化のシミュレーションを利用した Biohandwriting を提案する。既存の手書き文字の自動生成手法では、ランダムにパラメータを変動させて新しい文字画像を生成するた

め、不自然に変動した画像が得られる場合があった。提案手法では、生態シミュレーションに用いられる遺伝的アルゴリズムのアイデアを応用することで、自然に見える文字の特徴を残して新たな文字を生成できる。したがって、提案手法により不自然な文字の生成を抑えて、文字認識器の学習に効果的な文字画像を生成できると考えられる。図2に提案手法を利用した学習データ拡張のイメージを示す。既存の手法では、ランダムな変動により生成された文字データを学習データに加えていたが、提案手法によって生成される自然な文字データを学習データに加えることで、文字認識器の性能改善が期待できる。また、自動生成した文字データによる学習データセットの拡張は、人手による手書き文字を収集してデータセットを拡張する場合よりもコストを抑えることができる。

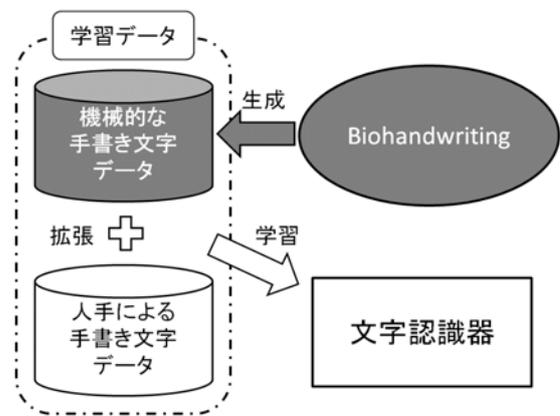


図2 提案手法を用いた学習データの拡張

3.2. 形態の進化シミュレーション

形態の進化シミュレーションとは、遺伝的アルゴリズムのアイデアを用いて生物の形態の進化や発現現象をシミュレートするものである[3]。例として、R.Dawkinsの「バイオモルフ(生物的形態)」が挙げられる。バイオモルフは、生物の個体発生の数学的なモデルを用いて多様な形態を表現する。図3にバイオモルフのモデルの例を示す。図3のモデルでは“F”が「枝」を示し、“[]”が枝の分岐，“+”と“-”が分岐の方向を示している。

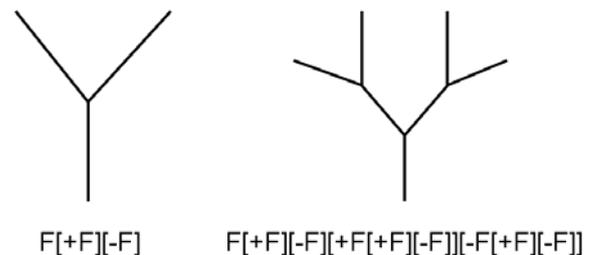


図3 バイオモルフのモデル

図3のようなモデルのパラメータに対して、遺伝的アルゴリズムのオペレータを適用することで、形態の遺伝

や個体の発生をシミュレートできる。遺伝的アルゴリズムのオペレータとは、「交差」や「突然変異」などの次世代の個体を生成する操作である。選択された親に対して遺伝的アルゴリズムのオペレータを適用することで、親の特性を受け継いだ個体を生成できる。図4に遺伝的アルゴリズムのオペレータの例を示す。遺伝的アルゴリズムでは、遺伝子を図4のような数値や文字の列で表現する。「交差」は、2つの遺伝子に対して特定の位置で値を組み替えて次世代の個体を生成する操作である。「突然変異」は、定めた確率で遺伝子の値をランダムに変動させて次世代の個体を生成する操作である。

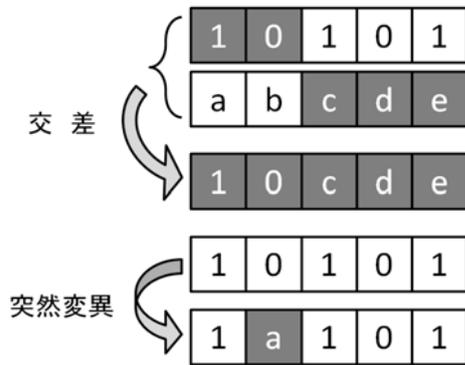


図4 遺伝的アルゴリズムのオペレータ

3.3. Biohandwriting

本研究では、形態の進化シミュレーションのアイデアを参考にし、文字の形状を表現するパラメータを変動させて自然な特徴をもつ文字を自動生成する。図5に文字の進化の例を示す。第1世代の左下と右上の文字を親として選び、新しく第2世代の文字を生成した。生成された第2世代の文字は、親の特徴を受け継いでいることがわかる。

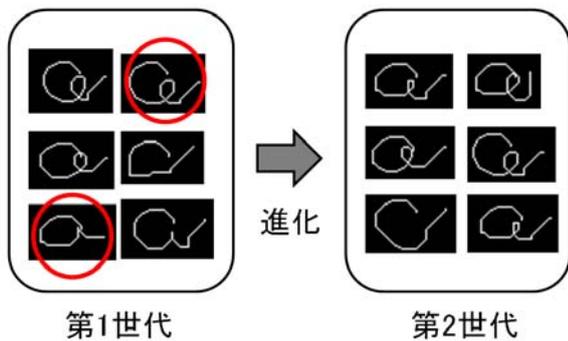


図5 進化の例

Biohandwriting における個体の遺伝子は、文字を構成するストロークの Delta-Log Normal Model のパラメータで表現する。図6に遺伝子の例を示す。図6に示すように、文字を表現する遺伝子は単一のストロークを表

す実数値の列で表現する。単一のストロークを表す実数値は、開始点の座標と Delta-Log Normal Model のパラメータの組に対応する。

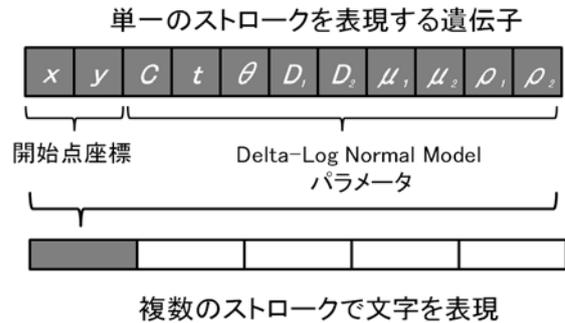


図6 文字を表現する遺伝子

親の選択では、自然に見える2つの個体を人手により選出する。選出された親の遺伝子に対し、交差と突然変異を適用して次世代の個体16個を生成する。交差の方法は1点交差として、交差の位置は毎回ランダムに選ぶ。また、突然変異が起こる確率は0.3とし、該当する遺伝子の値を平均0、標準偏差0.03の分布をもつ乱数で変動させた。以上の交差と突然変異に関するパラメータは予備実験の結果により定めた。

4. 評価実験

4.1. 評価実験の概要

Biohandwriting によって生成した文字画像と、テンプレートをランダムに変動させた文字画像について、形の「自然さ」と「ばらつき」の2つの観点から比較する。また、似た文字同士の特徴の類似性についても評価する。

本実験では、英小文字26文字を対象として、各文字につき2つのデータ集合 $A = \{\text{Biohandwriting による生成データ 50個}\}$ 、 $B = \{\text{ランダムな変動による生成データ 50個}\}$ を用いた。生成に必要なテンプレートは、文字のストロークを近似する Delta-Log Normal Model のパラメータである。また、データ集合Aは、Biohandwriting による5世代までの進化を3回試行して得られた文字240個から、無作為に抽出した文字50個で構成した。

実験1では、データ集合A、Bに含まれる不自然な文字の割合を比較する。生成された文字の自然さは、人手により判断した。

実験2では、データ集合A、Bの文字画像について7つの Hu Moments 不変量[4]を求めて、文字形状のばらつきを比較する。Hu Moments 不変量は、パターンや文字の位置や拡大・縮小、傾き、反転に依存しない特徴量である。したがって、Hu Moments 不変量の値で文字形状の特徴を表すことができる。

4.2. 実験1の結果

生成した文字の一部を図7に示す。図7(a)が提案手法により生成した文字、(b)がランダムな変動で生成した文字である。

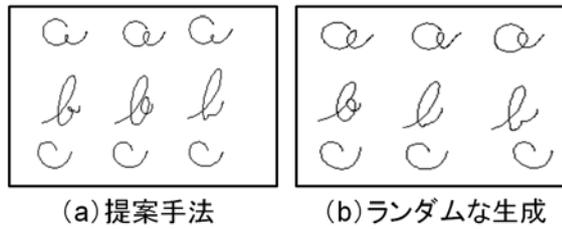


図7 生成した文字の例

自動生成した文字について、人手で不自然な文字を判別した結果を表1に示す。実験対象のすべての文字について、不自然に見える文字は、データ集合Aで約5%、データ集合Bで約23%含まれていた。したがって、提案手法では、不自然な文字の生成をランダムな変動による生成手法よりも約20%抑えることができた。

表1 不自然な文字の割合

データ集合	不自然な文字数 (総文字数)	割合
すべてのA	62 (1300)	約5%
すべてのB	295 (1300)	約23%

また、比較的文字を構成するストローク数が多い文字“k”や“m”では、提案手法とランダムな変動による生成手法の両方で、不自然な文字が30%以上含まれていた。文字を構成するストローク数が多いほど不自然な文字が生成される傾向にある。

4.3. 実験2の結果

例として、“a”、“b”、“c”、“m”、“w”について、生成された画像のHu Moments不変量の標準偏差を表2(a)、(b)に示す。

表2 Hu Moments 不変量の標準偏差

(a) データ集合Aの一部

	S1	S2	S3	S4	S5	S6	S7
a	0.366	0.309	0.710	0.158	0.077	0.111	0.203
b	0.319	0.710	0.987	0.229	0.447	0.273	0.305
c	0.238	0.133	0.273	0.128	0.103	0.081	0.110
m	0.526	2.560	8.567	1.706	19.855	4.822	23.993
w	0.272	0.939	1.383	0.844	2.846	1.514	0.432

(b) データ集合Bの一部

	S1	S2	S3	S4	S5	S6	S7
a	0.392	0.258	0.706	0.084	0.072	0.058	0.065
b	0.396	1.201	1.694	0.871	4.932	1.394	0.756
c	0.375	0.211	0.439	0.191	0.192	0.079	0.264
m	0.368	1.490	3.555	1.387	14.832	2.956	5.066
w	0.423	1.596	1.791	0.925	4.757	2.137	1.212

表2中の s_i ($i=1..7$)は、計測された7つのHu Moments不変量 h_i ($i=1..7$)の標準偏差である。データ集合Aとデータ集合Bの s_i を比較すると、多くの文字でデータ集合Bがデータ集合Aよりもばらついていることがわかった。す

べての s_i についてデータ集合Aがデータ集合Bより大きい値をとった文字は“g”、“i”、“m”の3文字だけであった。

また、似た文字同士として“a”と“d”、“m”と“w”、“p”と“q”のHu Moments不変量の平均値を比較すると、データ集合Aでは7つの S_i のうち6つが近い値をとり、データ集合Bにみられない特徴の類似性が確認できた。したがって、Biohandwritingによりランダムな変動では得られない特徴を捉えた文字を生成できたといえる。

5. 考察

提案手法により、不自然な文字の生成を約20%抑えることができた。また、ランダムな変動により生成した文字画像よりもばらつきを大きく損なった文字は3文字にとどまり、提案手法では自然で多様性のある文字を生成できたといえる。しかし、提案手法では単一の文字生成を対象としたため、連結した文字列画像の生成が困難であった。提案手法により生成した文字画像を手書き文字認識器の学習データとして利用するためには、連結した文字列を生成する工夫が必要である。また、提案手法によって生成された文字は、自然に見える文字の特徴を捉えているため、文字の自動生成に必要なテンプレートとして活用可能である。手書き文字の自動生成において、文字のテンプレートの多様性は重要な要素であり、自動生成した自然にみえる文字をテンプレートに追加することは、文字認識器の精度向上に有効である。

6. おわりに

手書きの文字認識器の学習データセット拡張を目指して、不自然な文字の生成を抑えたBiohandwritingを提案した。Biohandwritingでは、文字の変動に生物の形態の進化シミュレーションのアイデアを応用して、新たな文字画像を生成した。提案手法では、ランダムな変動による生成手法と比べて、不自然な文字の生成を約20%抑えることができた。今後の課題として、手書き文字認識器の学習データとして、連結した文字列を生成するための工夫が必要である。

参考文献

- [1] Tamás VARGA, Daniel KILCHHOFER and Horst BUNKE, “Template-based Synthetic Handwriting Generation for the Training of Recognition Systems,” Advances in Graphonomics: Proceedings of IGS, 2005.
- [2] Réjean PLAMONDON, Wacef GUERFALI, “The generation of handwriting with delta-lognormal synergies,” Biological Cybernetics 78, pp.119-132, 1998.
- [3] 伊庭 斉志, “遺伝的アルゴリズムの基礎,” オーム社, pp.198-209, 1994.
- [4] Ming-Kuei HU, “Visual Pattern Recognition by Moment Invariants,” IRE Transactions on Information Theory 8, pp.179-187, 1962.