

アフィン変換による正規化を用いた甲骨文字の認識率向上

Improving Recognition Rate of Oracle Bone Inscriptions Using Normalization by Affine Transformation

石井 康史†
Koji Ishii

藤川 佳之†
Yoshiyuki Fujikawa

孟 林‡
Lin Meng

山崎 勝弘‡
Katsuhiko Yamazaki

1. はじめに

甲骨文字は、3000 年以上前の中国殷代に使われた亀の甲羅や獣の骨に刻まれた象形文字であり、漢字の祖形とも言われている[1]。図 1 は甲骨文字の刻まれた甲骨の拓本[2]であり、拡大された文字は漢字の“貞”を表している。甲骨文字の解読は、我々の使う漢字の起源を知る上で非常に重要であるが、長い年月による劣化などが原因で認識しにくい問題がある。また、現在の解読は、歴史学者の経験や文脈と、従来からある文献によって行われることが多い。

甲骨の拓本の中には、図 1(b)のように傾いた文字や、テンプレートに比べて文字の縦横比が異なる文字が多数存在する。甲骨文字の認識に用いているテンプレートマッチングは、2 つの画像が似ているほど算出される類似度は高くなり 1 に近づくため、これらの文字は、テンプレートマッチングによってうまく認識することができない。そこで、本研究では、認識対象文字のアフィン変換による正規化を用いた認識精度向上を提案し、甲骨文字を正確かつ高速に認識することを目的としている。

これまで行ってきた甲骨文字の認識[3]~[5]では、まず、ガウシアンフィルタ、2 値化、ラベリングによってノイズを除去する。次に、細線化、ハフ変換によって文字の骨格を抽出する。最後にテンプレートマッチングによって骨格を抽出した画像とテンプレートを比較し、類似度を算出する。ここで用いるテンプレートは本学文学部で開発された、文字の傾きがなく、線分の太さと文字の大きさが統一された甲骨文字画像である[6]。

正規化では、文字の傾き、倍率、位置の変換が必要となるため、アフィン変換を用いて文字の正規化を行う。本研究では、アフィン変換によって、認識対象とテンプレートの画像内の文字領域を統一することでテンプレートマッチングによって算出される類似度の向上を目指す。

2 章では甲骨文字認識の流れを述べる。3 章では正規化の手法と正規化による文字の変化を説明し、4 章では正規化による実験結果を説明する。5 章でまとめと今後の課題を示す。

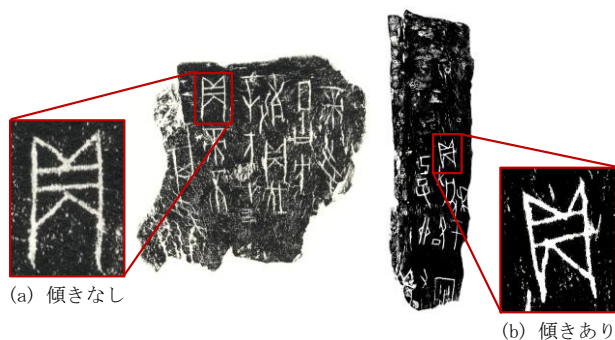


図 1. 甲骨に刻まれた甲骨文字

2. 甲骨文字認識の流れ

2.1 処理概要

図 2 に甲骨文字認識の流れを示す。原画像は、図 1 の拓本から切り出した認識対象となる甲骨文字である。まず、原画像に対して、ガウシアンフィルタと動的に閾値を決定した 2 値化によって、画像内の小さいノイズを除去する。次に、ラベリングを行い、動的に閾値を決定することによって、画像内の大きいノイズを除去する。さらに、細線化によって、文字の線幅を 1 に収縮し、ハフ変換によって文字の骨格である直線を抽出する。

最後に、テンプレートマッチングを用いてテンプレートとの類似度を算出することによって文字の認識を行う。ここで用いるテンプレートは、甲骨文字データベースより抽出した認識対象文字に対応するテンプレートである。なお、本研究では、認識の閾値を 0.4 としている。

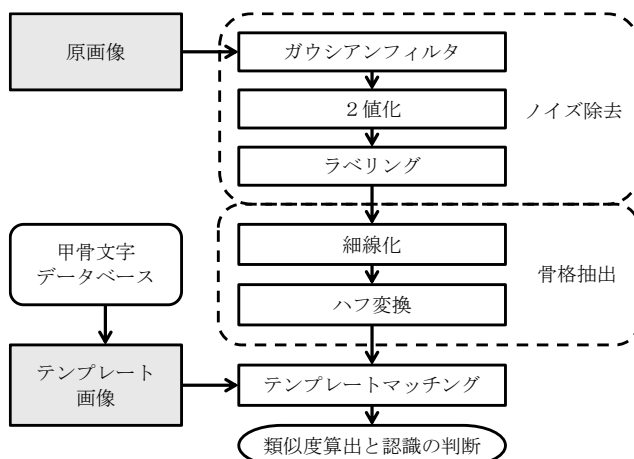


図 2. 甲骨文字認識の流れ

2.2 認識の過程

図 3 に甲骨文字認識の処理例を示す。

(1) ガウシアンフィルタ・2 値化

ガウシアンフィルタは、画像内の雑音を抑制するための平滑化に必要な処理である。式(1)のガウス関数に従って重みを変化させた加重マトリクスを用いて空間フィルタリングを行うことで、画像を平滑化する。

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cdots (1)$$

平滑化された画像に対して 2 値化を行うことで、画像内の小さいノイズの除去が可能である。2 値化は大津法によって閾値を動的に決定している。

†立命館大学大学院理工学研究科, Graduate School of Science and Engineering, Ritsumeikan University

‡立命館大学理工学部, College of Science and Engineering, Ritsumeikan University

(2) ラベリング

ラベリングは、2 値画像に含まれる各連結成分にラベル（番号）を割り当てる操作のことである。ラベリングによってラベル付けされた連結成分の中から閾値以下のものをノイズとして除去することで認識対象から大きいノイズの除去が可能である。

動的な閾値の設定には拡張した p-タイル法を用いており、画素数をソートして最も変化の大きいところを閾値として決定している。

(3) 細線化

細線化は、与えられた図形の線幅を細めて、幅 1 画素の中心線を抽出する処理のことである。本研究では、田村法を用いて細線化を行っている。

田村法は、3×3 のパターンを用意し、画像の走査から該当する画素を保存しておき、走査の後で該当画素の値を 0 にする。この処理を該当画素がなくなるまで繰り返すことで、文字の中心線を抽出することが可能である。

(4) ハフ変換

ハフ変換は、検出したい線の形があらかじめ決められており、代数方程式でその形状が表現できる場合に有効である。本研究では、文字の直線情報を検出するためにハフ変換を用いている。

直線を代数方程式は式(2)を用いる。

$$\rho = x \cos\theta + y \sin\theta \quad \dots \dots (2)$$

ここで、 ρ は座標原点から直線へ下ろした垂線の長さ、 θ は垂線と x 軸との間の角度を表すパラメータである。(2)式によって、x-y 空間中の直線は、 ρ - θ 空間中の 1 点として表される。 ρ - θ 空間中に描いた軌跡から、多数の軌跡が集中している位置を求めることで画像中の直線が検出できる。

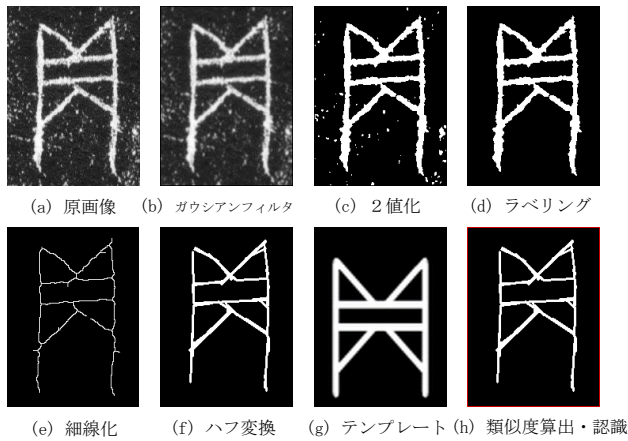


図 3. 甲骨文字認識の処理例

2.3 正規化の必要性

甲骨文字の中には、甲骨に沿って書かれたために傾いているものや、テンプレートに比べて文字の縦横比が異なるものが多数存在する。図 4(a)はその一例である。このような文字に対してテンプレートマッチングを行った場合、図 4(c)のように認識対象とテンプレートに差異が生じるため、期待される類似度を得ることができない。

そこで、認識対象の画像に含まれる文字領域をテンプレートに含まれる文字領域と一致するように変換（正規化）することでテンプレートマッチングによって算出される類似度を改善し、認識率の向上を目指す。



図 4. 認識率低下の要因

3. 正規化の手法

3.1 正規化の流れ

図 5 に正規化の流れを示す。正規化は、図 1 のハフ変換まで処理を行った画像に対して処理を行う。まず、文字の傾きを修正するために、回転角度を変化させて画像の回転を行う。次に、回転した画像に対して、テンプレートとの文字領域の大きさを統一する拡大縮小を行う。さらに、統一された文字領域を画像内の同じ座標に統一することで変換を完了する。最後にテンプレートマッチングによって、類似度が最大となった画像を正規化結果として出力する。



図 5. 正規化の流れ

3.2 アフィン変換による正規化

本研究では、文字の正規化に 2 次元のアフィン変換を用いる。アフィン変換は画像の回転、拡大縮小、平行移動を行うことができるため[7][8]、正規化に必要な条件を満たしている。アフィン変換の式を式(3)に示す。

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix} + v \quad \dots \dots (3)$$

ここで、 x' 、 y' は変換後の座標を示しており、T の行列の値によって回転、拡大縮小の処理が可能となり、v の行列の値によって平行移動が可能である。

3.2.1 文字の回転

画像の回転を行う場合、(3)式の T、v は式(4)のように設定する。これによって、2 次元図形を原点中心として θ だけ回転することが可能である。

$$T = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, v = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \dots \dots (4)$$

アフィン変換による角度の正規化は、図6のように認識対象の文字を一定角度ずつ回転し、出力されたすべての画像に対して以降の処理を行う。本研究では、甲骨に含まれる文字の傾きの傾向から、回転角度の範囲を-30度～30度に設定し、5度ずつ変化させて出力している。

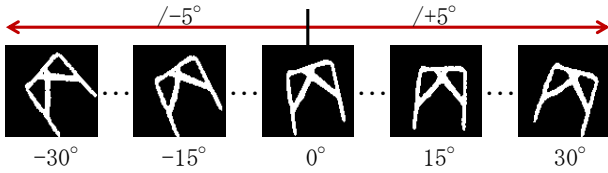


図6. 文字の回転

3.2.2 拡大縮小

回転を行うことで角度を正規化した画像は、図7のようにテンプレートに比べて文字領域の大きさが異なる。この文字領域の大きさを拡大縮小によって統一する。

拡大縮小を行う場合、(3)式の T 、 v は式(5)のように設定する。ここで、 S_x 、 S_y は x 方向、 y 方向それぞれに拡大縮小する倍率を表す。

拡大縮小に用いる倍率は、認識対象とテンプレートの文字領域の大きさを比較し算出する。

$$T = \begin{pmatrix} S_x & 0 \\ 0 & S_y \end{pmatrix}, v = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \cdots \cdots (5)$$



(a) 回転後 (b) テンプレート (c) 倍率の差異

図7. 文字の倍率の差異

3.2.3 位置調整

拡大縮小を行った画像は、図8のように、文字領域の大きさが統一されているため、平行移動によって2つの画像の文字領域を一致させることが可能である。

平行移動を行う場合、(3)式の T 、 v は式(6)のように設定する。ここで、 x_t 、 y_t が平行移動する値である。

$$T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, v = \begin{pmatrix} x_t \\ y_t \end{pmatrix} \cdots \cdots (6)$$



(a) 拡大縮小後 (b) テンプレート (c) 位置の差異

図8. 文字の位置の差異

3.3 テンプレートマッチング

図9に示すように、正規化によって、認識対象とテンプレートの文字領域を一致させることができたため、最後に回転によって出力されたすべての画像に対してテンプレートマッチングを行うことで、最も類似度の高い画像を最終的な正規化の結果として出力する。



(a) 正規化結果 (b) 正規化後の差異

図9. 正規化結果の類似度

テンプレートマッチングによって算出する類似度の関数は、正規化相互相関関数であり、式(7)で計算する。

$$S_{NCC}(d_x, d_y) = \frac{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} g(i,j)f(i,j)}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} g(i,j)^2} \sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} f(i,j)^2}} \quad (7)$$

ここで、 $g(i,j)$ 、 $f(i,j)$ は認識対象とテンプレートの画素の輝度値を示している。2つの画像が似ているほど、 $m \times n$ 次元空間上に投影した入力画像とテンプレートのそれぞれの直線からなる θ が小さくなり、内角 $\cos \theta$ は1に近づく。

4. 実験

4.1 実験内容

甲骨の拓本から抽出した 250 文字と、それぞれに対応するテンプレートを用いて正規化を行い、正規化の精度と類似度の向上を確認する。

本研究で用いる CPU は、Intel Core i7-4770 プロセッサで、動作周波数は 3.40GHz、64bit の命令セットで、OS は ubuntu14.04 である。プログラミング言語は、C 言語 (gcc バージョン 4.8.2) と OpenCV (バージョン 2.4.9) を用いており、ハフ変換のみ OpenCV で行って、それ以外の処理はすべて C 言語で行っている。

4.2 実験結果

表1に各回転角度における正規化結果の代表例を示す。原画像からハフ変換までの処理を行った画像に対して正規化を行った結果、テンプレートに比べて、文字の傾き、大きさ、位置が近づいていることがわかる。それに伴い、算出された類似度は正規化前に比べて大幅に増加している。

図10は、正規化前と正規化後の類似度の全体比較である。横軸は文字の番号であり、縦軸は類似度を示している。正規化前に比べて、類似度が向上した文字が全体の95%であり、平均2.15倍の向上率を達成した。

図11は、類似度が最大となったときの回転角度の文字数を示している。横軸は回転角度で、縦軸はその角度における文字数を示している。類似度が最大となる文字は、0度を中心に集中しており、-10度～10度の範囲で86%の文字の正規化が可能である。

表 1. 正規化による類似度の向上

回転角度	原画像	ハフ変換	正規化	テンプレート	類似度		
					正規化前	正規化後	向上率
0					0.311	0.553	1.78
5					0.162	0.501	3.09
10					0.192	0.620	3.23
-15					0.259	0.484	1.87
-20					0.169	0.565	3.34
-25					0.160	0.473	2.96
-30					0.039	0.706	18.10

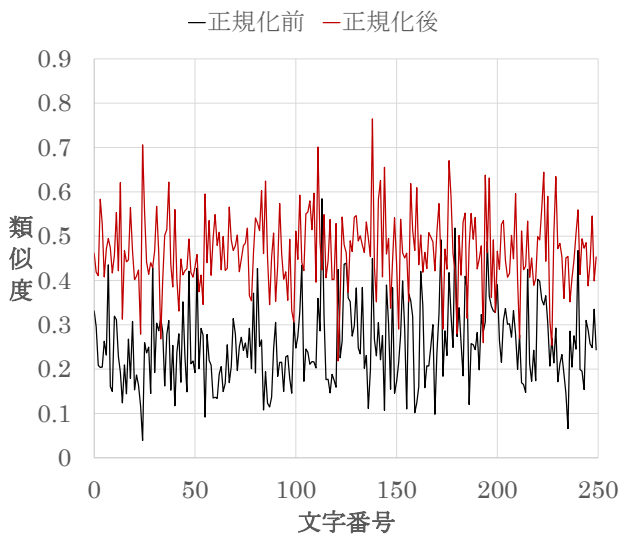


図 10. 正規化前後の類似度

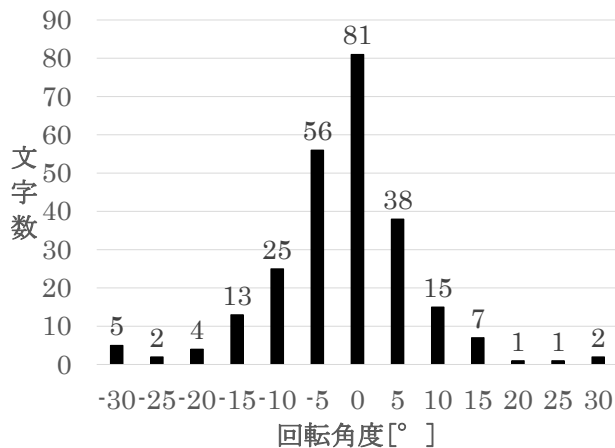


図 11. 類似度が最大となる回転角度と文字数

4.3 考察

表 1 から、文字の正規化は、テンプレートに合う最良の結果が得られている。正規化後の画像がテンプレートに比べて少し傾いているものもあるが、これは、回転角度の変化を細かく変化させることでさらに高い認識率を得ることができると考えられる。また、図 11 のグラフから、回転範囲を限定することで処理時間を短縮することができると考えられる。

図 10 では、甲骨文字 250 文字中、類似度が向上した文字は全体の 237 文字 (95%) であり、認識可能な文字は 214 文字 (86%) となったため、正規化は十分有効であるといえる。

また、複数の候補テンプレートを用いた場合でも、正確な類似度が得られるか実験が必要である。さらに、現在、1 文字の認識を CPU で行っているが、GPU による並列処理によって複数文字を同時に認識することも重要である。

5. おわりに

本研究では、甲骨文字の認識率向上を目指し、アフィン変換による文字の正規化を提案した。原画像からノイズを除去し、骨格を抽出した正規化対象となる画像に対して、アフィン変換による回転、拡大縮小、位置調整を組み合わせた正規化を行って、認識率の向上を達成した。

甲骨文字 250 文字に対して実験を行った結果、237 文字 (95%) の認識率が向上し、平均で 2.15 倍の向上率を達成した。また、認識できる文字は 214 文字 (86%) である。

今後の課題としては、複数の候補テンプレートの自動抽出とその正規化、GPU による超並列処理を用いた複数文字の同時認識などが挙げられる。

参考文献

- [1] 落合淳思, 甲骨文字に歴史を読む, 筑摩書房, 2008.
- [2] 濮茅左, 上海博物館蔵甲骨文字, 上海辞書出版社, 2009.
- [3] 石井康史, 藤川佳之, 孟林, 山崎勝弘, “甲骨文字認識における文字データベースの作成”, 電子情報通信学会総合大会, D-12-10, 2015.
- [4] 孟林, 石井康史, 藤川佳之, 落合淳思, 泉知論, 山崎勝弘, “甲骨文字認識プロジェクト—画像処理と GPU を用いた甲骨文字の高速認識”, 電子情報通信学会総合大会, D-12-11, 2015.
- [5] 藤川佳之, 孟林, 泉知論, 山崎勝弘, “ラベリングと 2 値化における動的な閾値設定を用いた甲骨文字の認識”, 第 77 回情報処理学会全国大会, 2P-01, 2015.
- [6] 落合淳思, 甲骨文字データベース
<http://koukotsu.sakura.ne.jp/top.html>
- [7] 長尾智晴, 安居院猛, 長橋宏, “遺伝的手法を用いた 2 値図形のパターンマッチング”, 電子情報通信学会論文誌, Vol. J76-D-II, pp.557-565, 1993.
- [8] A.Horimatsu, R.Niwa, M.Iwamura, K.Kise, S.Uchida, S.Omachi, “Affine Invariant Recognition of Characters by Progressive Pruning”, The Eighth IAPR International Workshop on Document Analysis System, pp.237-244, 2008.