

甲骨マッピングに関する研究

森脇 信啓[†] 孟 林[‡] 林 磊[‡] 安 成浩[§] 泉 知論[‡] 小柳 滋[†]

立命館大学 情報理工学部[†] 浙江大学 人文学院[§] 立命館大学 理工学部[‡]

1. はじめに

甲骨文字は 3000 年以上前に亀の甲羅、牛の骨などに刻まれた文字で、漢字の祖形とも言われる。これらの解読は文字の起源、歴史の研究において、非常に重要である。しかし、1899 年の初めての発見まで、長い年月が経過して、甲骨の劣化が進行し、多数の甲骨が断裂された。そして、分散された断片が増えて解読が困難になった。従って、甲骨マッピングにより、分散された大量の甲骨断片から、同一である甲骨を抽出し、統合するのが甲骨文の研究の重要な課題となる。

本研究では、画像処理[1]と確立統計の技術を用いて、甲骨の断片を分析し、甲骨のマッピングを行い、甲骨文の解読に貢献することを目指す。具体的には、甲骨の断片に対して、ラプラシアンフィルタにより画像のエッジを抽出する。そして、画像の収縮によりノイズを除去し、四辺からスキャンすることにより、甲骨の外郭を抽出する。さらに、甲骨の外郭の正規化を行う。最後に、パターンマッピング手法により、正規化された甲骨のマッピングを行う。

2. 甲骨断片と甲骨マッピングの先行研究

2.1 甲骨断片

図 1 は北京大學珍藏甲骨文字[2]に収録された甲骨断片の拓本を示している。各断片の下に甲骨の番号を示している。図 1 および[2]を見ると、規則のない甲骨断片が多数存在している。さらに、甲骨には亀裂や文字の刻みの跡などが残されている。これにより、甲骨マッピングが難しくなっている。

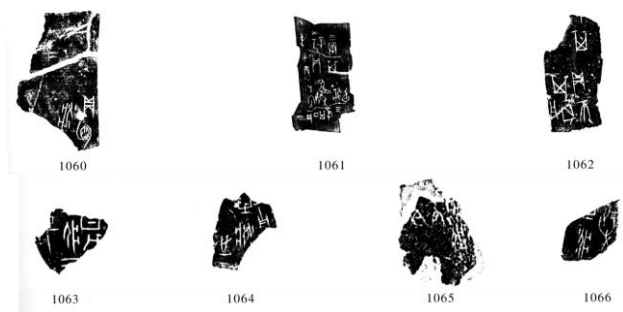


図 1. 甲骨の断片

2.2 先行研究

1973 年、周鴻翔らは初めて計算機を用いた甲骨マッピングの実験をアメリカで行った。しかし、当時の計算機性能の原因で、殆どのマッピングを手動で行った。

A Research about Oracle Bone Matching
Nobuhiro Moriwaki, Lin Meng, Lei Lin, Chenghao An,
Tomonori Izumi and Shigeru Oyanagi

[†] College of Information Science and Engineering, Ritsumeikan University

[‡] College of Science and Engineering, Ritsumeikan University

[§] School of Humanities, Zhejiang University

1977 年、童恩正らは、亀の甲羅に対して、甲骨マッピングを行った。この手法では、亀の甲羅の特徴を用いて甲羅を 9 分割し、マッピングの 6 個の条件（時代、字跡、骨盤、エッジ、内容）を使用して、マッピングを行った。しかし、この手法の成功率が僅か 40%しか満たされていなかった[3, 4]。

2001 年から、王愛民らが、甲骨を骨の種類に分けて、マッピングを行ってきた[5, 6]。しかし、亀の甲羅に対する研究が多い。

3. 甲骨文字認識の流れ

我々が考案した甲骨マッピングの流れは以下のようになっている。現時点では、(1)~(3)までを実現しており、本論文はそれらについて説明していく。

- (1) ラプラシアンフィルタによるエッジの抽出
- (2) 画像の収縮
- (3) 甲骨の外郭の抽出
- (4) 外郭の正規化
- (5) パターンマッピング

3.1 ラプラシアンフィルタによるエッジの抽出

ラプラシアンフィルタは空間 2 次微分を計算し、エッジを検出するフィルタである。注目画素の座標 (x, y) の画素値を $f(x, y)$ 、ラプラシアンフィルタにより求める新しい画素値を $L(x, y)$ とすると、求める画素値は以下のように表される。

$$L(x, y) = f(x-1, y-1) + f(x, y-1) + f(x+1, y-1) + f(x-1, y) + f(x, y) * (-8) + f(x+1, y) + f(x-1, y+1) + f(x, y+1) + f(x+1, y+1)$$

図 2(a)に元の甲骨文字画像、図 2(b)にラプラシアンフィルタの実行結果の 2 値化画像を示す。これよりエッジが抽出されたことが確認できる。

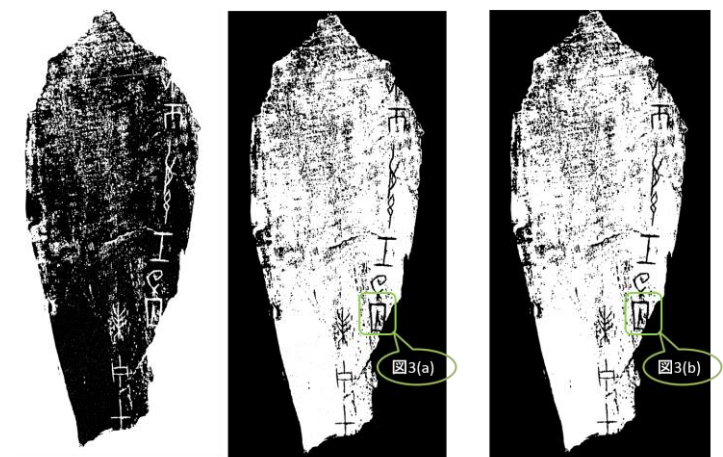


図 2. ラプラシアンフィルタと画像収縮の結果

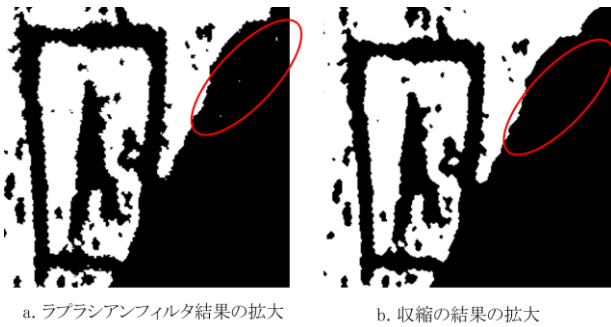


図3. ラプラシアンフィルタと収縮後の拡大

3.2 画像の収縮

ラプラシアンフィルタの結果を拡大して確認すると、いくつかのノイズが残っている。図3(a)では、赤い丸で囲まれている白点がノイズである。それを取り除くために、画像の収縮を行った。画像の収縮では、白い注目画素の8近傍の白い画素の数を数え、閾値以下になると、注目画素の反転を行う。

式を以下のように示す。 $C(x, y)$ が注目画素の8近傍の総和である。画素値 $f(x, y)$ が白い場合は、 $C(x, y)$ が閾値以下になると、画素値 $f(x, y)$ を黒くする。

$$C(x, y) = f(x-1, y-1) + f(x, y-1) + f(x+1, y-1) + f(x-1, y) + f(x+1, y) + f(x-1, y+1) + f(x, y+1) + f(x+1, y+1)$$

図2(c)は、画像収縮の結果である。図3(a)はラプラシアンフィルタ結果の一部の拡大図、図3(b)は収縮結果の一部の拡大図である。図3(a)の赤い丸には、白いノイズが存在している。図3(b)の赤い丸には、白いノイズが存在していないことにより、ノイズの除去が成功した。

3.3 甲骨の外郭の抽出

本研究では、画像の四辺から対面の辺へ、垂直方向で、白い画素が見つかるまで、スキャンする。白い画素が見つかる場所に白い画素として記憶し、その他の画素を黒くする。

図4を用いて、甲骨外郭検出の方法を説明する。図4(a)では、左辺から、右辺へ垂直方向で、白点の $f(x1, x1)$ が見つかるまでスキャンし、該当箇所を白くする。同様に、右辺から、左辺へ垂直方向で、白点の $f(x3, x3)$ が見つかるまでスキャンし、該当箇所を白くする。そして、その他の点は黒くする。これにより、図4(b)に $f(x1, x1)$, $f(x3, x3)$ が外郭となる。

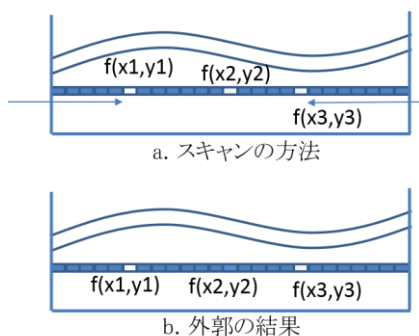


図4. 甲骨の外郭検出の方法

図5(a)が甲骨の外郭結果である。しかし、甲骨の特徴により、外郭を抽出できても、いくつかの課題が残っている。図5(b)は収縮結果の拡大図で、その外郭の拡大図は図5(c)で、これにより、甲骨の外郭の外に、小さい白い塊が存在している。さらに、図5(d)は収束結果の拡大図で、その外郭の拡大図は図5(e)で、文字の刻みにより、外郭がうまく閉じることができない。従って、外郭の抽出の更なる研究はこれからの重要な課題となる。

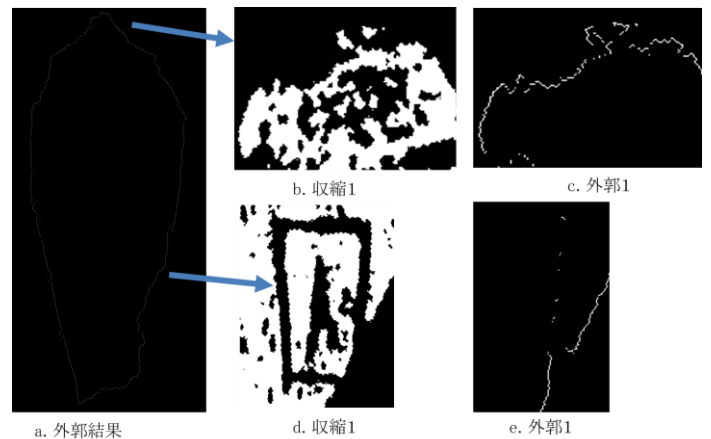


図5. 甲骨の外郭

4. おわりに

本稿では、甲骨文の解読を目指して、甲骨マッチングを行う。分散されている甲骨断片に対して、ラプラシアンフィルタにより画像のエッジを抽出する。そして、画像の収縮によりノイズを除去し、四辺をスキャンすることにより、甲骨の外郭を抽出する。さらに、甲骨の外郭の正規化を行う。最後に、パターンマッチング手法により、正規化された甲骨のマッチングを行う。

しかし、現時点では、外郭の抽出までしか実現していない。さらに、外郭の抽出の結果からみると、いくつかの問題が生じてしまう。今後、更なる甲骨外郭の抽出のアルゴリズムを更新し、外郭マッチングを行っていく予定である。

謝辞

本研究は、「文部科学省私立大学等経常費補助金」の助成により行われている。

参考文献

- [1] 田村秀行: コンピュータ画像処理, Ohmsha, 2004.
- [2] 李鍾淑, 葛英会編: 北京大学珍蔵甲骨文字, 北京大学震旦古代文明研究中心学術特刊, 上海古籍出版社, 2008.
- [3] 董恩正, 張陞楷, 陳景春: 關於使用電子計算機綜合商代卜甲碎片的初步報告, 考古 1977-3, pp.205-209, 1977.
- [4] 朱彥民: 殷墟考古發掘與甲骨文研究(上), 花木蘭文化出版社, 2012.
- [5] A.M.Wang, W.Y.Ge, Z.Zhao, Y.Q.Ge, G.Y.Liu and Q.S.Li: Research on computer matching of inscriptions on tortoise fragments, Computer Engineering and Design, Vol.32, No.10, pp.3570-3574, 2011.
- [6] S.Chen, A.M.Wang and T.W.Huang: Study on Strategic Approaches of Cultivating Tourism Market Entity in China, the 6th Int. Conf. on Computer Sciences and Convergence Information Technology (ICCIT), pp.342-347, 2011.