一种基于图形识别的甲骨文分类方法

X X X 1 X X X 1,2 X X X 1 X X 1 X X X 1,2

（1.X X X X X X X X X X X X X X X X XXXXXX）1

（2.X X X X X X X X X X X X X X X XXXXXX）2

**摘要:** 作为中国早期的象形文字，甲骨文的识别一直是一个难题。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。在小范围的实验中表明，这种方法能够较好地实现对甲骨文文字的自动分类。

**关键字：**曲率直方图，傅里叶描述子，甲骨文识别，甲骨文分类

A Classification Method for Oracle-bone Inscription

Based on Graphic Identification

XX XX-XXX1, XX XXXX-XXX1,2 , XXX XXX-XXX1, XXXX XXXX1, XXXX XXXX-XXX1,2

(X X X X X X, X X, X,XXXXXX) 1

(X X X X X X X, X X, X, XXXXXX) 2

**Abstract**  The Chinese oracle-bone inscription is one of the earliest known hieroglyphic inscriptions, and it has been a difficult problem for decades to automatically recognize the inscription. X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X . The experiment in the small scope indicates that the proposed method is translation, rotation and scaling invariant, and its performance presents that Chinese oracle-bone inscription can be classified correctly.

**Keywords** Curvature histogram，Fourier descriptor，Recognition of oracle-bone inscription，Classification of oracle-bone inscription

# 1 引言

早在三千五百多年的中国商代，甲骨文就已形成了成熟的文字系统，主要用于记录商王朝统治者的占卜活动。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。在这些甲骨上，反复出现的文字有4千多个，能够被专家辨识的约有2500多字。

虽然从文字的数量和结构上看，甲骨文已经发展成为比较严密的文字系统，形成了“象形、会意、形声、指事、转注、假借”等造字方法，但它毕竟处于早期汉字的成形阶段，带有明显的图画痕迹，不仅不具备现代汉字的标准笔划，在字形结构上也存在较大的随意性。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。如图1所示，具体分析如下：



图1 两个X X X X X X X X X X X X X X X X

（1）作为象形字，X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 写法。

（2）甲骨文需要用刀刻在硬骨甲壳上，因刀有锐有钝，骨质坚松有别，所以刻出的笔画粗细不一，有的纤细如发，有的浑厚粗重。X X X X X X X X X X X X X X X X X X X X X X 。所以整体上，甲骨文的笔画变化较其他文字复杂。

（3）为了 X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X借助。

基于X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 奏效。

# 2 相关研究

尽管X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 难点。

李锋[1] X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X ，如图2所示。

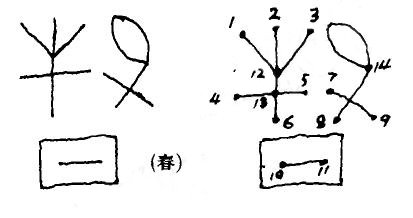


图2 根据X X X X X X X X X X X X X X XXX X

由于X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 准确性。

考虑到甲骨文本质上更接近图形符号，所以图形识别领域中基于轮廓线的特征描述的方法应有助于甲骨文的分类与识别。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。近期，A. El-ghazal[16]等又将多尺度曲率过零点的统计作为特征，其实验效果优于同期其他方法。

目前X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 自动分类上。

本文X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 方向。

# 3 曲率直方图及傅里叶描述子

## 3.1 基本思想

为获得X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 特征。

## 3.2 曲率的计算

由于X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 曲率。

具体X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 公式：

其中*，*X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 二阶导数。

## 3.3 曲率直方图的构建

为了X X X X X X X X X X X X X X X X X X X X X X X X X X 如下：

其中，X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X总个数。

然后X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 如图3：

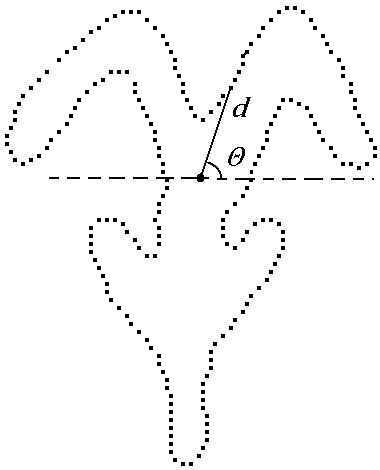


图3 计算采样点的与

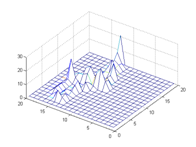
为了X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 量化：

其中，是直方图的大小，为所有采样点距重心距离的最大值。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。该过程可用如下公式表示：

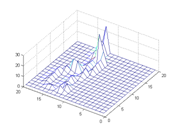
这里

其中X X X X X X X X X X X X X X X X X X X X X X X X X X X集合。

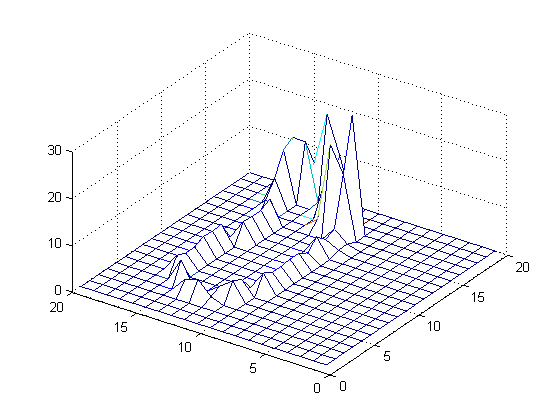
直方图X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 完整。



（a）



（b）



（c）

图4 曲率X X X X X X

如图4所示，X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 相似性。

## 3.4 傅里叶描述子

曲率直方图虽然能够较为准确、全面地反映图形的特点，但它本身并不具备旋转不变性和尺度不变性。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。为此，本文对曲率直方图进行了二维傅里叶变换，变换公式如下：

对获得的傅里叶系数进行取模，即可获得旋转不变性，再将所有傅里叶系数除以第一维的系数进行归一化，即可获得尺度不变性。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。归一化可通过如下公式完成：

最终得到X X X X X X X X X X X X X X X X X X X X X X X X X X X X 不变性。

# 4 分类方法的实现与实验结果

在完成X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 实验。

我们X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 测试。

我们选择本文提出的FDCH作为分类实验的特征，共125维。X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 。即80张测试图片中有72张分类正确。具体每类的分类结果见表5.2。

表5.2 X X X X X X X X X X X X X X

|  |  |  |
| --- | --- | --- |
| X X X | X X X X X | X X X X X |
| X | 2 | 80% |
| X | 0 | 100% |
| X | 1 | 90% |
| X | 0 | 100% |
| X | 1 | 90% |
| X | 0 | 100% |
| X | 2 | 80% |
| X | 2 | 80% |

# 5 总结与展望

实验X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 前景。

通过X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X 研究方向。

# 参考文献：

1. 李锋 等，甲骨文自动识别的图论方法，电子科学学刊，1996年12月第18卷 S1期
2. 周新伦 等，甲骨文计算机识别方法的研究，复旦学报（自然科学版），1996年10月，第35卷，第5期
3. Cheng-Lin Liu，In-Jung Kim，Jin H. Kim，“Model-based stroke extraction and matching for handwritten Chinese character recognition,” Pattern Recognition, Volume 34, Issue 12, December 2001, Pp 2339-2352
4. Rusinol，M. Llados. J. ，“Symbol Spotting in Technical Drawings Using Vectorial Signatures”, LECTURE NOTES IN COMPUTER SCIENCE, NUMB 3926, 2006, Pp 35-46.
5. Mickaël Coustaty, Stéphanie Guillas, Muriel Visani, Karell Bertet, Jean-Marc Ogier , “On the Joint Use of a Structural Signature and a Galois Lattice Classifier for Symbol Recognition”，Graphics Recognition. Recent Advances and New Opportunities: 7th International Workshop, GREC 2007，Pp: 61 – 70.
6. D. Zhang and G. Lu, “Review of shape representation and description techniques,” The Journal of the Pattern Recognation Society, vol. 37, no. 1, 2004.
7. S.X. Liao and M. Pawlak, “On image analysis by moments,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 3, 1996.
8. D. Zhang and G. Lu, “Generic fourier descriptor for shape-based image retrieval,” in Proc. of the IEEE ICME, 2002, vol. 1, pp. 425–428.
9. D. Zhang and M.C.Y. Lim, “An efficient and robust technique for region based shape representation and retrieval,” in IEEE/ACIS ICIS., 2007, pp. 801–806.
10. S. Tabbone, O. Ramos Terrades, and S. Barrat, “Histogram of radon transform. a useful descriptor for shape retrieval,” in Proc. of the IEEE ICPR, 2008, pp. 1–4.
11. D. S. Zhang and G. Lu, “Study and evaluation of different fourier methods for image retrieval,” International Journal of Computer Vision, vol. 23, no. 1, 2005.
    1. El-ghazal, O. Basir, and S. Belkasim, “A new shape signature for fourier descriptors,” in Proc. of the IEEE ICIP, 2007, pp. 161–164.
12. S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, 2002.
13. H. Ling，D.W. Jacobs，“Shape classification using the inner-distance,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 2, 2007.
14. C. T. Zahn，R. Z. Roskies，“Fourier descriptors for plane closed curves，” Transactions on Computers，vol.c-21, no. 3, 1972.
15. A. El-ghazal，O. Basir，S. Belkasim，“A novel curvature-based shape fourier descriptor，” in Proc. Of the IEEE ICIP, 2008, pp. 953–956.