

计算机科学

## 信函上英文地址的自动识别和翻译

屠晓<sup>1</sup>, 陈国跃<sup>2</sup>, 吕岳<sup>1</sup>

1. 华东师范大学 计算机科学与技术系, 上海 200062;

2. 秋田县立大学 电子与信息系, 日本秋田 015-0055

收稿日期 2007-9-13 修回日期 2008-2-3 网络版发布日期 2008-5-20 接受日期 2008-2-4

摘要

根据地址的语言特性, 定义和归纳了适用于邮政信函英文地址自动识别和翻译的用语规则. 针对由字符识别技术获得的信函地址, 提出了一种基于非精确字符串匹配技术的地址翻译方法, 自动识别出英文地址并翻译成中文. 实验结果验证了该方法的有效性, 能较好地减少识别错误带来的影响, 提高系统的翻译性能.

关键词 [非精确匹配](#) [基于规则](#) [地址识别](#) [地址翻译](#) [OCR](#)

分类号 [TP391.2](#)

## Automatic recognition and translation of english address on postal mails(Chinese)

TU Xiao<sup>1</sup>, CHEN Guo-yue<sup>2</sup>, LU Yue<sup>1</sup>

1. Department of Computer Science and Technology, East China Normal University, Shanghai 200062, China;

2. Department of Electronics and Information Systems, Akita Prefectural University, Akita 015-0055, Japan

### Abstract

According to the characteristics of the language used in address, rules different from those of natural languages were defined and applied to automatically recognize and translate English address on postal mails. To deal with the address got from Optical Character Recognition, an English-to-Chinese address translation method based on inexact string matching technology was proposed to recognize English address and translate it into Chinese. The experimental results showed that the present method is capable of reducing the influence of OCR error, and improving the translation performance.

**Key words** [inexact matching](#) [rule-based](#) [address recognition](#) [address translation](#) [OCR](#)

DOI:

通讯作者 吕岳 [ylu@cs.ecnu.edu.cn](mailto:ylu@cs.ecnu.edu.cn)

### 扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(763KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

相关信息

▶ [本刊中 包含“非精确匹配”的  
相关文章](#)

▶ 本文作者相关文章

· [屠晓](#)

· [陈国跃](#)

· [吕岳](#)