

ICS 35.240.20
L 67



中华人民共和国国家标准

GB/T 17961—2000

印刷体汉字识别系统要求与测试方法

Requirements and test methods for identify
system of printed Chinese character

2000-01-03 发布

2000-08-01 实施

国家质量技术监督局 发布

前 言

本标准是非键盘输入标准,它规定印刷体汉字识别系统的功能、技术要求、测试方法,该标准的制定和实施将正确引导印刷体汉字识别系统的研制与开发。

本标准的附录 A 是提示的附录。

本标准由中华人民共和国信息产业部提出。

本标准由中国电子技术标准化研究所归口。

本标准起草单位:“八六三”计算机智能主题专家组。

本标准主要起草人:顾小凤、丁晓青、刘迎建、钱跃良。

中华人民共和国国家标准

印刷体汉字识别系统要求与测试方法

GB/T 17961—2000

Requirements and test methods for identify system of printed Chinese character

1 范围

本标准规定了印刷体汉字识别系统的功能、技术要求和测试方法。
本标准适用于印刷体汉字识别系统,也适用于其他有关系统。

2 引用标准

下列标准所包含的条文,通过在本标准中引用而构成为本标准的条文。本标准出版时,所示版本均为有效。所有标准都会被修订,使用本标准的各方应探讨使用下列标准最新版本的可能性。

- GB 2312—1980 信息交换用汉字编码字符集 基本集
- GB/T 7589—1987 信息交换用汉字编码字符集 第二辅助集
- GB/T 7590—1987 信息交换用汉字编码字符集 第四辅助集
- GB/T 12345—1990 信息交换用汉字编码字符集 辅助集
- GB 13000.1—1993 信息技术 通用多八位编码字符集(UCS) 第1部分:体系结构
(idt ISO/IEC 10646.4:1993)
- GB 13131—1991 信息交换用汉字编码字符集 第三辅助集
- GB 13132—1991 信息交换用汉字编码字符集 第五辅助集

3 系统功能

3.1 基本功能

3.1.1 预处理

将普通纸媒体文本在自动或人工干预下通过光学输入设备将印刷信息转换成适合识别的灰度(含二值)图像。

3.1.2 识别

将上述图像所包括的印刷符号转换为 GB 2312、GB/T 7589、GB/T 7590、GB/T 12345、GB 13000.1、GB 13131 和 GB 13132 所规定的编码信息。

3.2 扩展功能

3.2.1 版面分析

应将版面自动分成块,并正确标明每个块的属性,对文字块还需标明各块之间连接关系的逻辑序号。

3.2.2 表格识别

应能正确识别表格线,并可将表格和文字建立对应关系。