

ICS 33.050.01

M 30

**YD**

# 中华人民共和国通信行业标准

YD/T 1890-2009

---

## 信息终端设备信息无障碍 辅助技术的要求和评测方法

Technical Requirements and Evaluation Methods for Assistive  
Technology of Information Accessibility in Information Terminals

2009-06-15 发布

2009-09-01 实施

---

中华人民共和国工业和信息化部 发布

## 目 次

前 言	II
引 言	III
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
4 辅助技术和辅助功能的需求分析和分类	4
4.1 信息交流障碍类型	4
4.2 针对身体障碍和习惯差异障碍的辅助技术和辅助功能	4
4.3 针对感官/感知障碍的辅助技术和辅助功能	5
4.4 针对认知障碍或文化差异的辅助技术和功能	5
4.5 针对沟通障碍的辅助技术和功能	6
4.6 针对混合障碍的辅助技术和功能	6
4.7 辅助技术分类	6
4.8 辅助功能分类	6
5 辅助技术要求	7
5.1 语音识别技术	7
5.2 生物识别技术	8
5.3 语音合成技术	10
5.4 语音放大技术	11
5.5 屏幕阅读技术	11
5.6 光学字符识别（OCR）技术	12
6 辅助功能要求	14
6.1 语音控制功能	14
6.2 声音转译功能	14
6.3 语音转换功能	14
6.4 视觉显示辅助功能	15
6.5 盲文显示功能	15
6.6 字幕功能	15
6.7 图示/图标功能	15
6.8 操作提示/反馈功能	15
7 评测方法	16
7.1 评测方法概述	16
7.2 主观评测方法	16
7.3 客观评价方法	19
附录A（资料性附录） 增强键盘的扩展功能	21

## 前 言

本标准第7.2节对应于ITU-T P.851《基于对话系统业务的主观评测》。

本标准与语音识别技术相关的术语与GB/T 21023-2007《中文语音识别系统通用技术规范》一致。

本标准的附录A为资料性附录。

本标准由中国通信标准化协会提出并归口。

本标准起草单位：工业和信息化部电信研究院

本标准起草人：孙悦、刘册、落红卫、朱晓峰、易高雄、张薇

## 引 言

信息无障碍环境的建立，不仅能够提高残障人士的生活质量，同时也方便健全人的生活。倡导和建立信息无障碍环境体现了重视人权平等的社会风气和人文关怀。提高人人参与机会的均等，积极持续地发展信息无障碍服务是和谐社会的一个具体表现。

建立信息无障碍环境，发展信息无障碍业务，设计、研发信息无障碍产品都需要技术标准的支持。通过技术标准的制定，可以：

- (1) 指导公共环境和服务的无障碍设计以及对信息无障碍环境和服务的评估；
- (2) 引导企业自主研发无障碍设备，对设备信息无障碍设计进行评估；
- (3) 带动企业实行产品创新，创造新的价值链；
- (4) 帮助政府立法、实施信息无障碍环境。

因此，信息无障碍标准的制定不但能够推动信息无障碍环境和服务的建立，同时还能促进相关产品的研发，从基础研究和人性设计上提升产品的应用价值。

# 信息终端设备信息无障碍辅助技术的要求和评测方法

## 1 范围

本标准分析了不同残障类型对辅助技术和辅助功能的需求，并对信息无障碍的辅助技术和辅助功能进行了分类，规定了目前主流信息无障碍辅助技术、辅助功能的技术要求和评测方法。

本标准适用于支持信息无障碍辅助技术和辅助功能的信息终端设备（包括公共设施、家用设备等）和业务。

## 2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准。然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

- GB 2312 信息交换用汉字编码字符集——基本集
- YD/T 789 免提电话机技术要求和测试方法
- YD/T 1538 数字移动终端音频性能技术要求及测试方法
- YD/T 1643 无线通信设备与助听器的兼容性要求和测试方法
- YD/T 1889 手柄电话助听器耦合技术要求及测试方法

## 3 术语、定义和缩略语

### 3.1 术语和定义

下列术语和定义适用于本标准。

#### 3.1.1

**辅助技术和辅助功能** Assistive Technology and Assistive Function

帮助使用人防止、补偿、减轻、制约任何应用障碍，使得使用人能够方便使用信息终端设备和业务的技术和功能。

#### 3.1.2

**信息无障碍** Information Accessibility

任何人（无论健全人还是有残障的人，无论小孩还是老年人）在任何情况下都能平等地、方便地、无障碍地获取信息、利用信息。

#### 3.1.3

**信息终端设备** Information Terminal Equipment

与用户有通信接口或应用界面，能够处理信息和/或支持通信功能的设备或系统。

#### 3.1.4

**词汇量** Vocabulary

语音识别系统所能处理的最大的词的个数。

3.1.5

**孤立字语音 Isolated Word Speech**

以字为发音单元，字与字之间需有有意、明显停顿的语音。

3.1.6

**连接词语音 Connected Word Speech**

以词为发音单元，词与词之间需有有意、明显停顿的语音。

3.1.7

**连续语音 Continuous Speech**

词与词之间无需刻意停顿的自然发音语音。包括朗读语音和自然语音。

3.1.8

**说话人自适应 Speaker Adaption**

语音识别系统通过用户发音完成对识别模型参数修改的自适应过程。

3.1.9

**环境自适应 Environment Adaption**

语音识别系统通过分析设备所处环境情况对识别模型参数修改的自适应过程。

3.1.10

**小词汇量 Small Vocabulary**

能识别的词处于 0~1023 个。

3.1.11

**中词汇量 Median Vocabulary**

能识别的词处于 1024~16383 个。

3.1.12

**大词汇量 Large Vocabulary**

能识别的词大于 16383 个。

3.1.13

**特定人 Speaker Dependent**

指语音识别系统的参数仅为某人特别建立。

3.1.14

**特定人群 Group Dependent**

指语音识别系统的参数为具有某种发音特性的群体建立。

3.1.15

**非特定人 Speaker Independent**

指语音识别系统可面向一般用户，不受说话人语音及口音限制。

3.1.16

**PC平台 PC Platform**

语音通过麦克风、PC上的声卡等输入设备完成语音采样并在PC上完成识别的系统。

3.1.17

**电话平台 Telephone Platform**

语音通过电话通道、声卡和传输并在相应的服务端完成识别的系统。

**3.1.18****嵌入式平台 Embedded Platform**

识别系统是嵌入在某个信息终端产品上（如HPC、PDA、手机和信息家电等），通过特定设备输入语音并完成识别的系统。

**3.1.19****分布式平台 Distributed Platform**

语音识别特征提取在前端嵌入式设备上完成，语音识别其他计算在后台系统和计算机上完成的系统。

**3.1.20****背景环境 Background Environment**

语音识别的语音采集设备所处环境中的噪声情况及待识别语音中的噪声情况。

**3.1.21****信噪比 Signal to Noise Ratio**

噪声情况以信噪比来衡量，信噪比定义为： $SNR=10 \lg (P_s/P_n)$ ，式中  $P_s$  定义为信号的功率； $P_n$  定义为噪声的功率。

**3.1.22****响应时间指标 Responding Time**

系统的响应时间统一采用实时系数概念。在系统的标准配置条件下，假设发音从  $T_s$  开始，发音结束时间为  $T_e$ ，识别结束时间为  $T_r$ ，则实时系数定义为：实时系数 =  $(T_r - T_s) / (T_e - T_s)$ 。

**3.1.23****识别字准确率 Word Correct Rate**

设正确文本字数为  $N$ ，识别结果文本字数为  $M$ ，按照识别结果文本与正确文本根据“最小代价匹配”原则运用动态规划算法，得到正确识别字数  $Mc$ 、删除错误字数  $D$ 、插入错误字数  $I$  和替换错误字数  $S$ ，有： $N = Mc + S + D$ ； $M = Mc + S + I$ ；

则识别字准确率  $WCR = ((Mc - I) / N) \times 100\%$

**3.1.24****指令检测错误率（等错率） Dectection Error Rate (Equal Error Rate)**

假设关键词表的词汇量为  $KW$ （个），检测语音长度为  $HR$ （小时），出现关键词  $N$ （次）， $C$  为每小时每个关键词最大容忍的误报个数（一般取 10），系统报出关键词  $M$ （个），其中，正确  $FD$ （个），错误  $FA$ （个）， $FD + FA = M$ 。则：

DET 曲线：以  $Far$  为横轴， $Frr$  为纵轴画出的曲线；

等错率  $EER$ ：DET 曲线上  $Far = Frr$  时， $Far$  或  $Frr$  的值。

**3.1.25****对话效率及对话正确响应率 Conversation Efficiency & Conversation Accurate Responding Rate**

对于口语语言对话类系统主要关心信息获取的效率，性能定义为：

$N$  = 系统完成一个信息查询并获取到正确信息所需的平均语轮数；

对话效率= (1/N) × 100%;

对话正确响应率=系统对信息咨询过程每次语音输入的平均正确响应率。

### 3.2 缩略语

下列缩略语适用于本标准。

ITE	Information Terminal Equipment	信息终端设备
OCR	Optical Character Recognition	光学字符识别
TTS	Text-To-Speech Engine	文字朗读引擎

## 4 辅助技术和辅助功能的需求分析和分类

### 4.1 信息交流障碍类型

信息交流障碍可以分为以下5类：身体和习惯差异障碍、感官感知障碍、认知和文化差异障碍、沟通障碍以及混合障碍。

总的来说，交流有障碍的人会因为身体问题或习惯差异不能进行自如的操作控制；因为感官感知问题不能获取设备信息；因为认知问题或文化差异障碍不能正确理解和操作设备；因为沟通问题不能明确表达自身需求，因此设备和业务在设计时，应充分考虑到对于不同的障碍类型，采用有针对性的辅助技术和辅助功能以帮助各类人群实现无障碍的信息交流。

### 4.2 针对身体障碍和习惯差异障碍的辅助技术和辅助功能

对于有身体障碍的人群，可以采用的辅助技术有语音识别技术、语音放大技术、语音转换技术、语音合成技术、语音控制技术。此外，可以提供除鼠标之外的多种信息输入、定位装置，如键盘、眼动仪等；特别的为了适应有习惯差异的人群，设备可以支持左、右手输入的方式等。并且应当尽量为这些输入提供足够大小及空间的按键以及避免基于小角度旋转的旋转控制。

表1给出了针对有身体障碍和习惯差异的人群可能使用到的辅助技术、辅助功能和辅助设备。

表1 针对有身体障碍人群提供的辅助技术、功能和设备

辅助的技术、功能、设备	说 明
语音识别技术	使手臂或手只能有限控制肌肉或完全丧失控制能力的人群可以通过语音方式输入文本或对用户接口实现控制
语音放大技术	通过放大语音信号，来满足听力下降人群的正常通信需求
语音转换技术	对于语言能力残障的用户，提供语音到文字、图像等方式的转换。例如失语的人群，可以借助那些有助于唇读和手语的设备和技术，如文字电话、视频电话、手语合成等进行交流
语音合成技术	对于口吃的人群，由于通信网络的回声和侧音问题会加重口吃的影响，同时由于网络负载的影响，一些弱音也并不能通过简单的放大被完全的补偿，因此可以采用语音识别技术和语音合成技术来改善
语音控制功能	对于灵巧型障碍和身体灵活性障碍的人群来说，可以通过语音控制功能来减少对设备按键、控制旋钮的操作，或减少身体的移动
免提功能和技术	力量/耐力障碍的人群，很难实现长时间的抓握，因此，电话的免提功能在这方面是有帮助的
辅助定位/点设备	针对手臂或手只能有限控制肌肉或完全丧失控制能力的人群，使他们可以控制使用鼠标的移动和功能。例如包括用脚操作的鼠标、用嘴控制鼠标的定位设备以及眼动跟踪系统等
辅助输入装置	对于灵巧性障碍的人群，使用鼠标很困难，很容易点击到屏幕上错误的部分而发生误操作，因此设备需要支持使用键盘等替代的输入信息方式(对于键盘的扩展功能，可以参考附录A)
屏幕键盘/触摸屏	针对灵巧性障碍及操作能力障碍的人群，屏幕键盘在屏幕上提供物理键盘的按键及功能，经常与辅助定点设备配合使用



表1 (续)

辅助的技术、功能、设备	说 明
预测型(联想式)字典	预测型字典根据使用者打出的字符做出单词预测从而加快了打字速度, 并且提供预测出的单词列表供使用者选择
左右手操作转换功能	有利于像左手操作等与大多数人不同习惯的使用者方便输入

#### 4.3 针对感官/感知障碍的辅助技术和辅助功能

对于有感官、感知障碍的人群, 所采用的辅助技术应能帮助他们使用感官感知以外的方式正常的获取设备信息, 实现人机交互。

表2给出了针对有感官/感知障碍人群的辅助技术和辅助功能。

表2 对于有感官、感知障碍人群提供的辅助技术、功能和设备

辅助的技术、功能、设备	说 明
视觉/听觉显示功能	为视力障碍人群提供屏幕放大显示(包括图形、窗口及文本)、为色盲人群提供颜色转变的显示、对于听力障碍人群提供图像、文字显示, 对于按键实现语音输出确认等
屏幕阅读技术	当通过语音输出设备或盲文显示器接收到文本或有限的图形信息时, 使视觉残障用户可以通过窗口、菜单以及控制实现巡航功能
盲文显示功能	对屏幕上显示的文本, 通过使用一连串针击转成盲文符号的形式, 提供逐行的盲文显示。有视觉障碍人群能够通过导航功能持续使用
文本转换功能	对于有视觉障碍的人群, 通过语音合成技术将电子文本转换为语音
振动或灯光显示功能	有听觉障碍的人群在听告警音、呼叫进程音及电话振铃时可能会觉得困难, 应采用提供的多种模式的信号表示方式, 如闪烁灯光、振动感知等
感应耦合技术	对于使用助听装置的人群采用线圈耦合以提高收到语音的清晰度
专用通信设备	例如文本电话, 是为有听力或语言障碍的人设计的电话。该话机上配备了键盘和显示屏。可以使用户发送和接收用键盘书写的信息。提供了一种使用户在电话线上使用文本终端进行交流的手段
语音转换	对于有听觉障碍的人群, 提供语音到文字、图像等方式的转换, 借助那些有助于唇读和手语的设备和技术, 如文字电话、视频电话、手语合成等进行交流
声音转译功能	针对有听觉障碍的人群, 提供声音信息的视觉转译。如系统beep声通过屏幕闪烁或类似的方式显示出来, 视频及静态图像通过隐藏字幕或相关技术加以描述
字幕功能	针对有听觉障碍的人群, 对视频媒体上的口头资料提供文本转译。包括重要计算机应用软件、远程学习、CD-ROM、视频会议及其他形式的交互式视频提供字幕信息
语音控制功能	对于那些触觉灵敏度缺失的人群, 尤其是具有假肢的人, 他们可能是不能使用电脑的触摸屏或触控板的, 因此要使用语音控制功能

#### 4.4 针对认知障碍或文化差异的辅助技术和功能

对于有认知障碍或文化差异的人群, 所采用的辅助技术和辅助功能应能帮助他们理解设备指示及操作方法, 实现人机交互。

表3给出了针对认知障碍人群的辅助技术和辅助功能。

表3 对于认知障碍或文化差异人群的辅助技术和辅助功能

辅助的技术、功能	说 明
符号界面功能	在文本方式外尽可能更多使用符号的方式进行相互作用和沟通。使认知障碍人群使用标记符号式文字和使用尽可能短和简单的说明
简化语言复杂性功能	提供一种可供选择的表述内容的方式, 这种方式具有复杂度较低的词汇量及表述形式, 以便于认知障碍人群理解
语音识别/认可	可以提供一种替代打字录入文本的工作方式, 如果识别字典可将某些使用者发出的声音调整到一个简化的范围
操作提示功能	对于复杂的操作流程或难于记忆的操作步骤, 提供下一步该进行何种操作的提示信息
多文字/语言支持功能	帮助来自不同文化区域的人群方便使用设备

#### 4.5 针对沟通障碍的辅助技术和功能

沟通障碍，主要表现在诵读困难及语言障碍上，从而不能明确地表明自己的信息。

对于沟有通障碍的人群，所采用的辅助技术和功能应能帮助他们准确表达自己的信息，实现人机交互。

表4给出了针对沟通障碍人群的辅助技术和提供的功能。

表4 对于有沟通障碍人群的辅助技术

辅助技术	说明
语音识别/认可	识别字典可将某些使用者发出的声音调整到一个简化的范围，从而识别出用户信息
语音转换	提供语音到文字、图像方式的转换。用简化的图形、符号表明用户信息

#### 4.6 针对混合障碍的辅助技术和功能

由于混合残障是指用户群具备两种及以上的残障问题，因此在辅助技术和功能的选择上要结合各种残障类型的辅助技术多方面考虑。

#### 4.7 辅助技术分类

综合上述对不同障碍形式的分析，辅助技术分为以下类别：

a) 语音识别技术：通过拾取语音信号，能够在一定时间内，以一定正确率识别语音语义或语音关键词的技术；

b) 生物识别技术：基于人的个体特征或行为进行身份判别的技术；

c) 语音合成技术：是将不清晰、有损伤的语音信号或者电子文本信息在一定时间内，以一定正确率合成为标准的语音信号的技术；

d) 语音放大技术：将语音信号（本标准主要涉及窄带音频信号，即 300~3400Hz 频带范围内的音频信号）在一定条件下进行放大的技术；

e) 屏幕阅读技术：信息终端（如计算机）上将用户的操作和/或选项以语音形式播报出来的技术。

f) OCR 识别技术：将纸质文本的内容扫描、识别为电子文本信息的技术；

g) 感应耦合技术：与助听器配合使用，使得佩戴助听器者接听电话时达到一定质量保障的技术。具体技术要求应符合 YD/T1889《手柄电话助听器耦合技术要求及测试方法》和 YD/T1643。

本标准分析了目前主流的辅助技术，随着技术的发展，可能会出现更多的辅助技术，本标准将及时修订或补充。

#### 4.8 辅助功能分类

综合上述对不同障碍形式的分析，辅助功能分为以下类别：

a) 语音控制功能：是指在语音识别的基础上，将语音命令转化为控制、操作信息终端的功能。

b) 免提/头戴功能：在语音通信中，不使用手柄可以直接实现语音对话的功能。免提技术具体的技术要求应符合 YD/T789 和 YD/T1538。

c) 声音转译功能：将常规以声音信号输出转化为光或振动信号输出的功能。

d) 语音转换功能：语音转换的基础是语音识别技术，在识别了语音的基础上，将识别的信息转换为文本、图像、图形、符号、动画、合成手语等方式的功能。

e) 视觉/听觉显示辅助功能：实现例如放大屏幕上的信息，转换内容颜色，亮度/对比度调节等功能；

f) 盲文显示功能：对输出的信息转换为盲文的功能。

g) 字幕功能：对音视频媒体上的语音信息提供文本转译。

h) 符号/图形界面功能：使用符号的方式进行相互作用和沟通。

i) 操作提示/反馈功能：对于复杂的操作流程或难于记忆的操作步骤，提供下一步该进行何种操作的提示信息。

j) 多文字/语言支持功能：根据设备具体应用场景，用户界面、操作提示等信息支持多文字/语言。

k) 左右手操作转换功能：方便左、右手操作的功能。

## 5 辅助技术要求

### 5.1 语音识别技术

#### 5.1.1 语音识别技术概述

语音识别是用复杂的软件技术和相关的硬件设施完成对人类的声音的辨别，而且把这些转换成对话指令，画面选择或数据选项等。

语音识别系统的应用场景主要包括以下3类：

a) 基本文字输入类：是指需要根据用户发出的语音，一字不漏的记录语音中所有的内容的应用场景。具体应用系统包括听写机、语音输入等。

b) 命令控制检测类：是指根据用户语音中发出的指令，完成特定操作和业务的应用场景。具体应用系统包括：语音控制、语音拨号等。

c) 口语语言对话类：是指接受用户以对话形式发出的自然的口头语言，明白及理解用户意图及想要获取的信息，并将以各种形式与用户进行反馈，以将对话继续进行的的应用。具体应用系统包括：口语对话系统、信息查询系统、电话订票系统等。

#### 5.1.2 不同场景下的技术要求

##### 5.1.2.1 对于基本文字输入类的环境要求及指标建议

- a) 应用环境：PC平台或嵌入式平台；
- b) 背景环境：信噪比达到15dB及以上；
- c) 词汇量：大词汇量；
- d) 应用人群：特定人或者非特定人；
- e) 工作模式：有限或无限朗读语音识别；
- f) 响应时间指标：实时系数1.5以下；
- g) 性能指标：识别字准确率达到85%以上。

##### 5.1.2.2 对于命令控制类的环境要求及指标建议

- a) 应用环境：电话平台或嵌入式平台；
- b) 背景环境：信噪比达到15dB及以上；
- c) 词汇量：小词汇量及以上；
- d) 应用人群：特定人群或非特定人；
- e) 工作模式：连接词语音识别或有限口语语音识别；
- f) 响应时间指标：实时系数1.5以下；
- g) 性能指标：指令检测错误率（等错率）小于30%。

##### 5.1.2.3 对于口语语言对话类的环境要求及指标建议

- a) 应用环境：PC平台、电话平台、嵌入式平台或分布式平台；

- b) 背景环境：信噪比达到15dB及以上；
- c) 词汇量：中词汇量及以上；
- d) 应用人群：非特定人；
- e) 工作模式：有限或无限口语语音识别；
- f) 响应时间指标：实时系数1.5以下；
- g) 性能指标：对话效率大于或等于40%；对话正确响应率应大于或等于75%。

5.1.3 语音识别系统的应用建议

- a) 在设备允许的位置范围内，识别系统应能保证滤掉背景噪声并使呼吸声降到最低；
- b) 为避免口音差别造成的错误识别，识别系统应能通过练习、校准的方式适应不同口音使用者使用；
- c) 允许超过一个字或词有相同的响应，即指令输入选择多样化（“多重映射”）；
- d) 选择语音为输入词汇、词组使最小可能的识别混乱；
- e) 语音识别系统可采用创建系统识别窗口，也就是训练使用者在每个输入之后暂停，选出字或字串判别；
- f) 为了使最小预测对话有效，应确保识别视窗启动时间反应灵敏而且期间充分；
- g) 提供在每个输入后的识别反馈，如重复操作或用键盘等其他方式确认；
- h) 如感觉声音反馈太慢，应考虑提供比较好的视觉反馈；
- i) 应给使用者提供辨别、修改来自任何不正确输入的机会；
- j) 给富有经验的使用者提供轮流指令优先级的可能性，如果使用者没有说某事或提供一个不适当的指令则只提供指令画面；
- k) 给富有经验的使用者提供使用捷径指令的可能性，避免必须经过冗长的对话；
- l) 考虑到特殊使用者可能在对话时不具有识别某种反馈的能力，因此，反馈应提供多种形式可供使用者选择；
- m) 对于不适合语音输入系统的人群，系统也应该提供其他方法告诉用户如何正确的退出该系统；
- n) 语音识别系统应有很好的人机界面，特别应考虑不同需求人群使用，提供“帮助文件”和多种提示方式给使用者使用。

5.2 生物识别技术

5.2.1 生物识别技术概述及一般建议

生物统计学识别技术是一种基于人的个体特征或行为进行判别的技术。传统的个人识别系统往往是通过密码的方式，如使用者遗忘或被人冒用将带来难以想象的后果。而生物统计学鉴别是基于个体特征的，能够避免上述问题。生物统计学技术包含面部成像（光学的和红外光），手和手指形状，以眼圈为基础的方法（虹膜和网膜），签名，声音，血管几何学，手指和指纹图像，这些技术的比较见表5。

表5 生物识别技术

生物体鉴别学	普及性	惟一性	持久性	可收集性	性能	可接受性	防欺骗性
脸	高	低	中	高	低	高	低
指纹	中	高	高	中	高	中	高
手型	中	中	中	高	中	中	中
虹膜	高	高	高	中	高	低	高
视网膜扫描	高	高	中	低	高	低	高

表 5 (续)

生物体鉴别学	普及性	惟一性	持久性	可收集性	性能	可接受性	防欺骗性
签名	低	低	低	高	低	高	低
语音签名识别	中	低	低	中	低	高	低
面部温度	高	高	低	高	中	高	高

一般的, 建议:

- a) 设备获取一个给定的生物统计特征应该是容易和舒服的;
- b) 数据采集和测量 (不包括触摸方式) 应使使用者感觉到亲和性;
- c) 生物统计学技术需要来自使用者的配合, 这种配合应让使用者觉得方便;
- d) 生物统计识别技术可能会涉及到个人隐私, 需让使用者清楚该项技术涉及哪些方面, 并自愿接受;
- e) 生物统计学识别系统推荐使用的人群是那些容易忘记或遗失身份证明的人群;
- f) 鉴别系统应保证是一个低错误率的系统;
- g) 对于典型的访问控制应用 (例如一个 ATM 机), 系统应该给出一个实时的判定;
- h) 应给用户 提供清晰、明确的使用说明书。

### 5.2.2 虹膜识别

生物统计学虹膜识别技术是通过使用照相机取得一个虹膜图像, 经与数据库中已有指定样本进行比对并加以判别的生物统计识别技术。在生命的最初两年虹膜的视觉材质是稳定的, 而且它的复杂纹理结构对于个体鉴别是很有用的信息, 虹膜鉴别是最准确的生物统计学鉴别方式。虹膜结构在出生后几年内趋于稳定, 如果没有经受外部伤害或白内障等疾病, 虹膜在人一生中基本保持不变。

建议:

- a) 采用虹膜识别技术的信息认证设备应具有“用户友好”的采集界面, 避免在数据采集时对使用者造成损伤;
- b) 识别过程应提供相应的语音提示或声音反馈;
- c) 虹膜识别系统应考虑一些特殊人群 (非常高的或非常矮的或那些坐轮椅的) 的使用便利性;
- d) 虹膜识别对于盲人可能存在不适用的情况。

虹膜识别的技术要求:

- a) 识别正确率: 虹膜识别正确率与硬件设备的视频分辨率、识别采样点的多少、系统算法复杂度等因素紧密相关。同时也与使用者的使用习惯相关, 经过训练后, 识别正确率应在95%以上;
- b) 识别响应时间: 虹膜识别响应时间与系统硬件处理能力、识别采样点的多少、系统算法复杂度等因素紧密相关。同时也与使用者的使用习惯相关, 应保证系统的识别响应时间为使用者所接受。一般应小于1S。

### 5.2.3 指纹识别

指纹是指手指尖内侧的生理图案, 它在胎儿时期形成并不会随成长改变的。每个人的指纹都是独特的, 甚至同卵双生儿的指纹都是不相同。指纹识别技术是指通过指纹扫描器取得指纹图像而且使用复杂的运算法则把图像转换成一个独特“映像”或分析图案, 与数据库中已有指定样本进行比对并加以判别的生物统计识别技术。

建议:

- a) 采用指纹识别技术的信息认证设备放置手指位置的表面应该是容易清理的;

b) 应有其他备用鉴别方式（如密码确认等），保证在使用者因为存在遗传基因、老化、环境或职业原因等因素造成误识别的情况仍可使用系统；

c) 识别过程应提供相应的语音提示或声音反馈。

指纹识别的技术要求。

a) 识别正确率：指纹识别正确率与硬件设备的视频分辨率、识别采样点的多少、系统算法复杂度等因素紧密相关。同时也与使用者的使用习惯相关，经过训练后，识别正确率应在95%以上。

b) 识别响应时间：指纹识别响应时间与系统硬件处理能力、识别采样点的多少、系统算法复杂度等因素紧密相关。同时也与使用者的使用习惯相关，应保证系统的识别响应时间为使用者所接受。一般应小于1s。

### 5.3 语音合成技术

语音合成技术是将文本状态的文字信息、预录语音或模板语音转化为可听的声音信息。在具体实现中，语音合成技术要在自然语言处理指导下，可以根据当前文本的特点，使用与之相适合的风格以及语气语调来完成文本到语音的转换。

语音合成技术要求可以应用在桌面系统和嵌入式系统中，这样用户在使用信息服务的时候，可以使互联网信息、短消息、电子邮件等多数以文本方式提供的信息可以用语音的方式输出。结合模式识别技术可以适用于更多场合。

通用技术指标有：

a) 可懂度：指语音中音节以上的语言单位（如字、词和句等）的可懂程度。人们在交谈时，并非将每一个字每一个音节均咬得很清楚，在语气、声调乃至情绪等方面都有很大的差异。人们在听人讲话时也是以听懂其意为目的，对个别字的发音并不在意。

b) 舒适自然度：自然度是指语音经信号变换后所复原的声音听起来和原来的语音在音色上的相似程度。自然度不理想，具体地说，就是节奏感较差，缺少轻重缓急和抑扬顿挫，不能为公众所接受。故此，恰当的韵律切分，正确地把握话语的韵律结构，是加强合成语音节奏感、提高其自然度的关键。自然度目前只能定性表示。

通用功能要求有：

a) 速度调整：通常要求语音合成设备或软件可以有控制手段可以根据使用者的实际要求来具体调节语音合成播放速度；

b) 音量调整：可以根据使用者的要求可以调节语音合成音量大小；

c) 发音风格：由于不同使用者喜欢不同的发音风格，所以在语音合成中最好可以有不同的发音风格来适应不同的使用者；

d) 通过对话管理（dialogue manager）提供输入信息的类型和数量，例如有注解的文本、正字法、焦点信息或作诗法信息等；

e) 响应发声的策略（例如通用语法或是简单的模板）、适应性（例如预定义的词汇或公开的词汇）、消息结构（语法，词汇选择）等；

f) 系统语音的数量，性别，是否是专业人员，是否培训过，韵律感等；

g) 语言：单语言或多语言合成，语音辨别能力等；

h) 语音合成类型：预录消息、模板语音、文本阅读、概念阅读。

## 5.4 语音放大技术

语音放大技术是指将语音信号进行放大，来满足使用者在收听或发送方向上的音量需求。

### 5.4.1 频率响应差

指额定工作条件下，300~3400Hz频率范围内各频率点的输出电压与规定频率点（1kHz）的输出电压的差值，以分贝表示。

一般，频率响应差应 $\leq 2\text{dB}$ 。

注：在某些情况下，特别地针对某些听力损伤患者，可以通过附加控制装置来增加可调频率整形功能。此外，频率整形还能用来补偿听筒耳承与人耳耦合时常有的低频声泄漏以及由于本地用户线路过长引起的高频跌落。

### 5.4.2 总谐波失真

总谐波失真是指额定工作条件下，测得的输出信号中总的谐波失真与总的输出电压的有效值之比。

在所有音量调节范围内，总谐波失真应 $\leq 2\text{dB}$ 。

### 5.4.3 信噪比

在一定额定条件下，在300~3400Hz频率范围内输出信号与各种噪声分量之和的比值，以分贝表示。

在所有音量调节范围内，信噪比应至少 $\geq 45\text{dB}$ 。

## 5.5 屏幕阅读技术

屏幕阅读技术是将信息终端上的可视化信息转化成语音信息输出的技术。

建议：

1) 屏幕信息可随屏幕上的光标指示文字、图形、符号等信息转化成语音输出，图形、图标应内嵌相应的文字信息。

2) 屏幕阅读技术应支持键盘“方位键”操作，支持焦点事件查询方式。

3) 光标移动速度和语音输出响应时间、阅读速度、音量等应根据使用者的需求进行调节。

4) 当光标移动到屏幕边缘应配有相应的提示音。

5) 应支持自定义快捷键方式，便于使用者对屏幕目标的选取。

6) 系统应配有语音导航、语音帮助信息便于使用者操作。

7) 具有该功能的手持通信终端设备（如手机），在收到信息或进行相应屏幕菜单操作时，应通过按相应功能键开启此功能。

8) 屏幕阅读技术应具备支持符合微软 SAPI 标准的语音引擎外挂的功能，可由用户自定义选择外挂不同的语音引擎（TTS）。

9) 语音音量可调节。

10) 朗读速度可调节。

11) 语调可调节（可选）。

12) 遇到英文单词，可选择分读（读字母）或连读（读单词）。

13) 遇到数字，可选择分读（读数字）或是连读（读数值）。

14) 支持中英文混读。

15) 能自动识别朗读简繁体汉字。

16) 按键反应时间可设置。

17) 具备复读功能。

18) 可选择是否朗读标点。

19) 朗读过程可控制，可以随时暂停、继续或停止。

20) 可为用户解释读出文字的字意，辅助理解多音字的现象。

21) 具备缓冲区复制功能，可将任何朗读出的信息复制到剪贴板。

22) 可以对 Windows 通用控件进行朗读，如桌面图标、任务栏、窗口的标题栏、菜单栏及菜单内容、地址栏、列表、状态栏的文字信息，能读出对话框、进度栏，各种组合框、可编辑框、单选按钮、复选框，各种按钮的名称和状态以及控制滑块，能够朗读出剪贴板上的文字。

23) 能朗读出鼠标指针处的文字信息，可读出图标和按钮的浮动条。

24) 支持事件即时朗读：当计算机产生某一事件时，比如焦点事件（选中、单击、右键等）、对话框弹出事件、菜单弹出事件、光标移动事件等，用户能够即时得到语音提示，从而采取进一步操作。

25) 屏幕朗读的工作模式，采用焦点方式取源或鼠标方式取源，并且二者可实现自由切换。

26) 支持微软公司的办公自动化软件。包括 Microsoft Word、Excel、PowerPoint、Access、写字板、记事本等。支持实时屏幕朗读，包括编辑光标位置、当前行内容、当前段内容、当前文档属性、光标移动过的内容。支持全文朗读、逐行逐字朗读，可朗读光标位置及光标所在处字体名称字号、朗读选中内容、朗读光标所在页数以及文章的总页数、总段数、总句数、总词数和总字数等功能。在朗读的同时，还支持点字显示（盲文输出）的功能（可选）。

27) 支持网页阅读，支持多框架结构网页，可以朗读网页全文，可运用网页综述功能获取当前页面的框架、链接、表单、表格、图片的数量，可选择逐行、逐句、逐词、逐字朗读，并可以任意选择网页内容，遇到特殊的焦点（如链接等），会听到特殊的提示。

28) 在浏览某一页面时，朗读到某一处，可按快捷键将该位置记下，当页面最小化后，切换到其他窗口操作，想继续浏览该最小化网页时，直接切换到该窗口，再按快捷键，可以从已标记处朗读。可以选择上翻页和下翻页以及跳到网页首和跳到网页尾的功能，从而快速浏览网页。可直接按快捷键切换到表单并可继续按此快捷键在多个表单之间切换。也可跳过链接或表单而直接朗读正文内容。按光标键切换到网页中的表格时，还可听读该单元格在当前表格中的行数和列数。在较熟悉的网页中，可“搜索”式朗读，通过查找关键字而直接跳转到需要朗读的内容处。切换链接时，遇到没有文字提示的链接（如图形链接），读屏会自动将该链接加序号，以方便区别。Outlook Express、Microsoft Outlook、Chm 格式文件和某些 EXE 格式电子书、帮助文档及其他包含网页格式应同样支持这些功能。

29) 实现各切换功能的所使用的快捷键可以由用户自定义、常用功能的快捷键可由用户自定义。如果系统自己提供默认的快捷键设置，则要注意这些设置在定义时应避免快捷方式之间的冲突，比如不能和微软操作系统的快捷方式冲突、还有联想、金山等软件的快捷方式、inter 显卡自己的快捷方式、笔记本电脑 F1-F11 的功能。

30) 支持点字显示器输出连接。可将屏幕取词结果时时翻译，并将翻译结果转换成点字显示器能够识别的 ASCII 码输送给点字显示器。用户可以通过触摸点字显示器上的盲文，来进行电脑的操作（可选）。

31) 支持 PDF 文档朗读（可选）。

32) 屏幕阅读技术的响应时间与系统硬件处理能力、系统算法复杂度、语料数据库大小等因素紧密相关。应保证系统的识别时间为使用者所接受。

## 5.6 光学字符识别（OCR）技术



### 5.6.1 概述

光学字符识别（OCR）技术是通过扫描仪等光学输入方式将各种票据、报刊、书籍、文稿及其他印刷品的文字转化为图像信息，再利用文字识别技术将图像信息转化为可以使用的计算机输入技术。

建议：

a) OCR 识别技术应具有可以自动判断、拆分、识别和还原各种通用型印刷体表格，并在表格理解基础上给出使用者预期的结果；

b) OCR 识别技术应能够自动分析文稿的版面布局，完成自动分栏，并判断出标题、横栏、图像、表格等相应属性，判定识别顺序；

c) OCR 识别技术应能将识别结果还原成与扫描文稿的版面布局一致的新文本；

d) 具有表格自动录入功能的 OCR 技术应能自动识别特定表格的印刷或打印汉字、字母、数字，并按表格格式输出；

e) 具有表格自动录入功能的 OCR 技术应支持将表格识别直接还原成 PTF、PDF、HTML 等格式文档，并可以对图像嵌入横排文本和竖排文本、表格文本进行自动版面分析。

本标准光学字符识别主要关注汉字识别。汉字识别系统通常分为手写汉字识别系统和印刷汉字识别系统两大类，如图 1 所示。

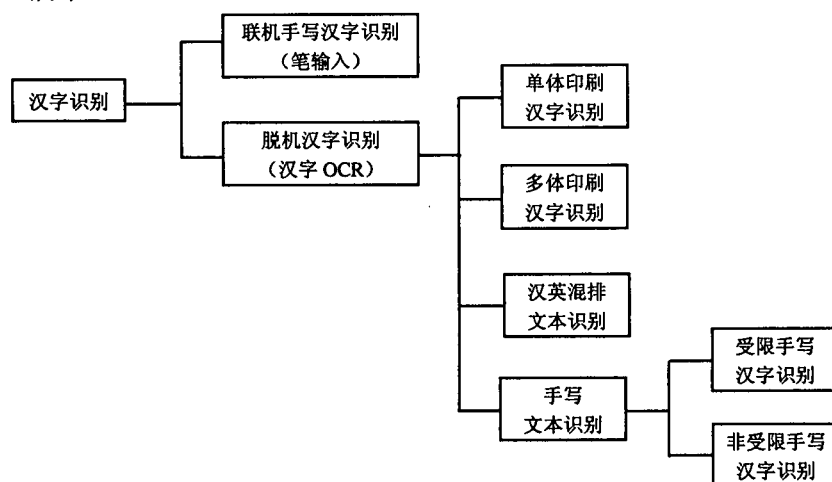


图1 汉字识别系统分类

按输入方式不同，手写汉字识别按输入方式不同又可分为联机和脱机两类。联机手写汉字识别一般采用一块专用的与计算机连接的书写板，人在书写板上书写字符时，字符信号即时直接输入计算机。这种方法也叫做笔输入方式，是一种实时输入的方法。脱机手写字符系统用来识别写在纸张上的字符，它同印刷光学字符识别一样，需用光电转换器（如扫描仪或摄像机）把字符图形变为电信号，再输入计算机进行识别。通常印刷字符的字形比较规范，识别比较容易，而手写字符的字形变化大，有的字符甚至人眼也很难直接辨识，使用计算机识别就更加困难，因此，通常以印刷汉字为测试样本来进行测试。

### 5.6.2 识别的字符类总数

系统识别的字符类总数决定整个系统的识别容量。

汉字识别系统应至少可以识别 GB 2312 规定的一级汉字，也可以适当增加一些较常用的二级汉字和专用汉字。具体字符集的规模和内容，可根据实际应用的需要仔细选用。

汉字识别系统的识别字符集应由汉字、标点、符号、数字、英文（其他外文）文字母等部分组成，字符总数应在 4000 之上。

### 5.6.3 识别的字体

只能适应单一种字体的印刷汉字识别系统为单字体印刷汉字识别系统。这种系统采用切换特征字典的方法来识别不同字体的印刷汉字。能用同一特征字典识别多种字体的识别系统为多字体印刷汉字识别系统。在设计多字体印刷汉字识别系统时，至少应识别宋体、黑体、仿宋和楷体这 4 种主要字体。

### 5.6.4 识别的字号

汉字识别系统应能识别目前所有字号共 16 种。对于出现在同一文本中的不同字号的汉字，在识别时应进行归一化处理。

### 5.6.5 识别率

识别率是系统最重要的指标。识别率分为单字识别率和系统识别率两种。单字识别率是指被正确切分的汉字图像被正确识别的概率，系统识别率是指识别结果相对原始文本而言的正确识别的概率。

注：从原始文本到识别结果输出，中间要经过扫描输入、行切割、字切割，然后才进行单字识别，中间各步都有可能影响识别率，尤其是“切割”这一步容易发生切分错误，造成字符的误识。

汉字识别系统在正常识别条件下的单字识别率应达到 98% 以上，系统识别率应达到 95% 以上。

### 5.6.6 识别速度

汉字识别速度分单字识别速度和系统识别速度两种。单字识别速度是指已被正确切分的汉字图像在单位时间内正确识别的总数，系统识别速度是指单位时间内正确识别的原始文本中的汉字总数。

## 6 辅助功能要求

### 6.1 语音控制功能

建议：

- a) 语音控制系统的响应时间应具有可调节功能，适应不同使用者的需求；
- b) 语音控制系统应提供指令自定义功能；
- c) 语音控制系统应提供训练、校准的方式保证使用者的指令成功率；
- d) 语音控制系统应提供可视化或语音导航等多种帮助信息方式供使用者选择；
- e) 语音控制系统应提供备用控制指令方式供使用者选择，在语音指令失灵时自动起用。

### 6.2 声音转译功能

声音信号是设计用来对使用者提供设备状态的信息。声音信号包括：装备警报信号（例如错误时的“beep”声）、忙音、传真自动应答音等。对于听力有障碍的人，建议对上述信号以光信号或振动信号等来代替。

建议：

- a) 对于不同的声音信号应能通过光信号的颜色、强弱和振动的强度、频率加以区分；
- b) 用户应可以根据自己的需要自定义不同声音信号与不同光信号、振动信号的对应关系。

### 6.3 语音转换功能

语音转换是在语音识别的基础上，将识别的语音信息转化为图像、图形、符号、动画等可视信息的技术。

建议：

a) 语音转换系统应保证转换后的图像质量、动画质量为使用者接受，并包含足够的语音信息。如语音转化成手语动画信息。

b) 语音转换系统应提供用户自定义信息形式功能，满足使用者将语音信息转换成自己所需形式的可视信息。

c) 语音转换系统应具有反应速度调节功能。

#### 6.4 视觉显示辅助功能

具有视觉显示辅助功能的系统应能实现以下功能：

a) 更改屏幕上文字的字体大小，更改屏幕上项目的大小；

b) 将放大更改为高对比模式，反转放大窗口的颜色；

c) 设定放大倍数，更改放大窗口的位置和窗口大小。

#### 6.5 盲文显示功能

点字显示器是一种盲文输出设备。

建议：

a) 该设备应具有盲文显示的点击功能，应既支持屏幕信息的盲文输出，同时支持输入字符的触觉输出确认；

b) 该设备应具有标准的通信接口满足与个人计算机相连接；

c) 该设备输出的盲文字符应符合盲文语法规则；

d) 该设备的软件字库应支持中文输出。

#### 6.6 字幕功能

具有该功能的系统在显示视频/音频和带有隐含文字说明的图像文件时应提供开启字幕显示的功能。

a) 字幕信息应在屏幕下方，减小对画面的影响；

b) 字幕显示速度应与播放的视频/音频画面同步，字体、字符大小应根据使用者需求进行调节。

#### 6.7 图示/图标功能

所有设备上的图示和图标在特定环境下用于标记控制键或者用来替代文字作为选项，或在屏幕上弹出通知窗口。

#### 6.8 操作提示/反馈功能

a) 应有多种操作提示方式供不同需求的使用者选择，如文字提示、图形提示、声音提示等。详见表6的输出信息分类及输出需求；

b) 应确保操作提示的便利性、准确性和有效性；

c) 特殊操作应有相应的警告提示，提示形式应采用醒目的提示窗口或采用特殊的声音；

d) 为盲人或视觉有障碍的使用者提供服务的信息终端，键盘输入，DTMF 拨号，应提供清晰的语音提示和反馈；

e) 为盲人或视觉有障碍的使用者提供服务的信息终端，在需要输入中文词语时，应提供清晰的字意解释，保证词语输入的正确性；

f) 使用者在进行密码输入等保密信息时，声音提示应采取规避处理，以保护客户信息的安全性；

g) 如果在打电话通话状态下，键盘输入，DTMF 拨号，应提供清楚语音提示和反馈。

表6 输出信息分类

通知的类型	输出需求
状态	显示需要显示继续状态，使用非干扰的光学的信号
警告	要求使用者注意音调，灯闪光，或提供在中心位置文本消息
反馈（响应）	提供动作的迅速承认，例如听得见的音调，灯闪光，显示改变，触觉，例如单击键的动作
数字	使用数字顺序的数字显示，考虑如何显示改变速率，例如拨号和钟面，和使用图表或表格
系统提示	系统需要从使用者得到输入/回答。使用可听见的信号立即获得使用者注意。使用特殊或可视化的信息指导使用者。把可视化的消息放置到使用者的视线中心。允许使用者不再需要消息可以取消消息。集成显示/控制以提供软键接口或提示使用功能键
文本	为文本消息使用小面板或语音输出。 考虑 CRT 或平面面板显示，为了文本文件使用足够大的文字数字的显示。需要与感人的易读性和安全已接受的标准相一致
图示	使用有迅速显示比率的高质量图像显示器来显示数字化可视化的图像

## 7 评测方法

### 7.1 评测方法概述

本标准对信息无障碍辅助技术和功能的评测主要包括两部分：一是主观评测，二是客观评测。两种评测方式都是从使用者感知的角度进行的。其中，主观评测方法适用于所有本标准提到的辅助技术和辅助功能；不同辅助技术的客观评测方法不尽相同，将分条描述。

### 7.2 主观评测方法

#### 7.2.1 概述

在进行主观评测之前，应对系统进行设置，使其系统尽量工作在独立工作模式，无需人为辅助或干预。在某个外部模块缺失或者某种设备因素使得系统无法独立工作时，可以考虑介入人为辅助，在不超出外部缺失模块的能力或者设备功能范围的前提下，尽量模拟缺失的模块功能或设备功能，并在此基础上得出近似的系统评测结果。

a) 在主观测试时应考虑系统所处环境因素，尽量贴近系统的真实使用环境：

1) 对于语音识别系统，应考虑设备所处环境的背景噪声，空间混响条件，语音采集设备本身的频率响应、信号传输方式带来的频谱影响等因素；

2) 对于语音合成的系统，应考虑设备所处环境的环境噪声，空间混响条件，语音播放设备本身的频率响应等因素；

3) 对于虹膜识别、OCR 和指纹识别的系统，应考虑设备所处环境的光照情况以及空气湿度和采集设备清洁程度、采集设备图像分辨率等各种因素的影响。

b) 在主观测试时也需要选择合适的测试者，考虑系统投放应用后可能接触到的各种使用者、尽量贴近系统设计针对的人群范围：

1) 对于语音识别系统应考虑主观测试者口音、方言、母语种类、年龄层次、性别、发音障碍等方面因素；

2) 对于语音合成系统应考虑选择主观测试者母语种类、年龄层次、听觉障碍、听力曲线等因素；

3) 对于虹膜识别系统应考虑测试者眼部疾病情况的影响、左右眼虹膜对称情况、隐形眼镜佩戴与否等因素；

4) 指纹识别则需考虑主观测试者指纹缺失或划伤、指部疾病情况、指部畸形、指纹稳定性的影响。

c) 在主观测试时也应考虑测试场景的选择问题，考虑测试内容是否超出系统设计所针对的范围，并且测试手段是否符合正常系统使用场景。

1) 对于语音识别系统应注意输入语音的词汇量是否超出系统限定、选取的语音内容的混淆度情况、说话速率快慢、语音发音强度的因素的影响；

2) 对于语音合成系统需考虑的因素有待合成内容的规范程度、合成内容是否符合正常语言习惯等因素影响；

3) 虹膜识别、OCR 及指纹识别则应注意采集图像时测试者的配合程度、采集设备与物体的对准程度、图像采集时的持续时间长度以及稳定程度等因素的影响。

d) 在主观测试时也应考虑到系统在投入应用后在人们的日常生活中的使用频率对最终主观评测结果带来的影响。应避免选取对待测系统具备相当经验的使用者或者对系统运作原理相当了解的专家人士进行主观评测，以避免测试者具备的先验知识引导测试者，使其操作让系统的运行效率超出一般情况，对评定结果带来的影响。对于使用频率相对较高和相对复杂的系统，应考虑系统对使用者的使用培训过程，尽量在使用者对系统功能有基本了解之后进行测试，给出中肯的结果。

7.2.2 主观评测方法

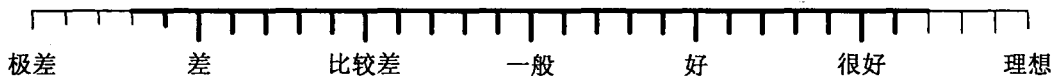
主观评测是从用户感知的角度进行，包括验证功能完备性和人机交互友好程度两部分。可采用主观评分制度进行划分，主要考察以下因素：

- a) 设备使用的便利性、易用性；
- b) 设备界面的友好程度、功能完善性；
- c) 提示、帮助信息准确性。

问题内容如下：

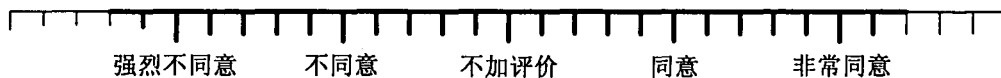
总体印象：

1. 对整个系统的总体映像。

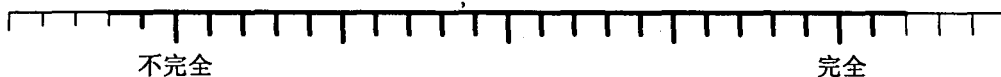


从系统获得的信息：

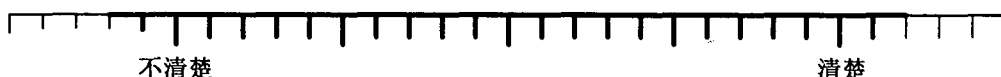
2. 系统提供了需要的信息。



3. 提供的信息..



4. 信息本身...



5. 请为提供的信息做综合评价。



与系统的交互:

6. 你对系统的理解程度。



7. 是否需要集中注意力、特殊的身体姿态来使系统识别用户信息?



8. 系统对环境是否过于敏感?



系统行为:

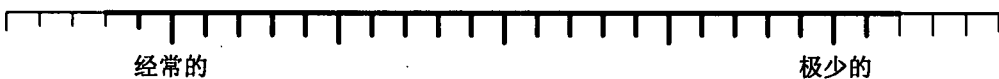
9. 在你的印象中, 系统正确地处理了你的指令?



10. 系统表现正常?



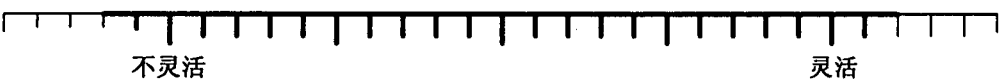
11. 系统是否经常出错?



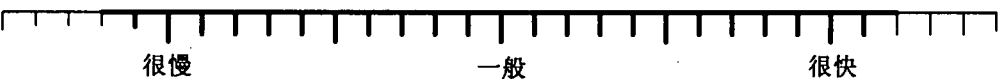
12. 系统表现的比较人性化?



13. 系统表现得灵活性?



14. 系统反应时间。



## 7.3 客观评价方法

### 7.3.1 语音识别评测

#### 7.3.1.1 基于语音标准库的测试

为保证语音识别系统测试的可重复性，测试应尽量采用基于语音识别标准库的测试方法。将标准库中预先录制好的语音输入待测系统，并统计系统输出结果。

待测语音识别系统应提供直接测试和间接测试两种接口。

**直接测试：**指利用被测系统带有的应用程序输入/输出接口，直接利用语音识别标准库中的语音文件进行测试。

**间接测试：**指测试系统利用高保真回放设备把语音识别标准库中的语音通过双方认可的方式输出到被测系统中。

#### 7.3.1.2 基于现场发音的测试

针对对于无法预先准备标准语音的系统，可以采取发音者现场发音输入语音识别系统的方法。测试基本情况同基于语音标准库的间接测试。

a) 需要有两个以上识别结果记录者，记录被测系统对当前发音的输出结果；记录表应包括发音人、记录人、操作人、监督人、发音内容、语音识别结果等内容；

b) 被测系统按照发音人应给出文件形式的输出结果，全部发音者测试结束后，统一按照性能标准进行指标评估，评估至少有两个人以上参与；

c) 给出与标准库测试相同的测试报告。

#### 7.3.1.3 客观评测的测试报告

报告应由以下几部分构成：

a) 被测系统的完整属性描述。

b) 测试数据的语音属性；测试词汇以及测试说话人的选择及确定情况。

c) 语音识别系统输出结果的统计；每个人识别各项指标以及平均识别指标。

d) 测试过程的情况纪录，采用的测试方法及运行过程的流畅性。

e) 被测系统的配置情况。

### 7.3.2 生物特征识别评测

测试配置如图2所示。

**第一步** 对采集的图像样本进行评估，虹膜图像考察指标：图像信噪比、虹膜图像有效面积、中心偏离度；要求虹膜识别系统的终端设备具有对采集后的图像输出能力；指纹考察指标：图像信噪比、指纹图像有效面积、指纹图像形变程度、中心偏离度；要求指纹识别系统的终端设备具有对采集后的图像输出能力。

**第二步** 评估过程应考虑到应用环境不同，对识别正确率有相应的影响，对设备应用环境的评估指标包括：环境照度、对比度的评估。

**第三步** 该图像样本进入被测试设备进行判别（图像数据库应采用国家相关部门发布的标准数据库）。

**第四步** 记录判别结果和响应时间。

**第五步** 取不同的测试图像样本多次测试，记录判别结果和响应时间。

第六步 计算识别正确率和平均识别响应时间。

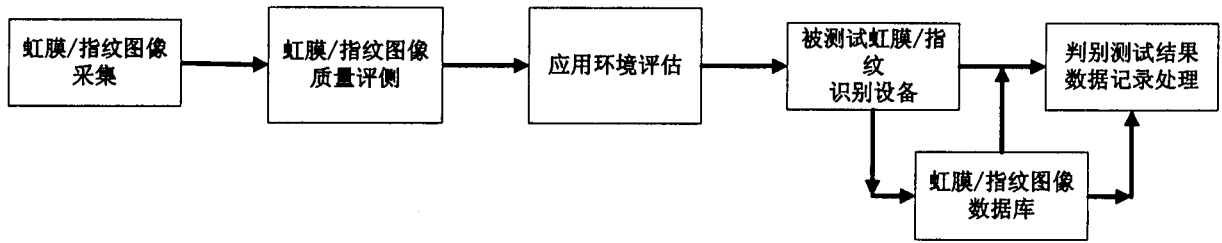


图2 虹膜识别设备测试配置

7.3.3 光学字符识别 (OCR) 技术评测方法

光学字符识别技术的客观评测是建立在原始样本和识别结果的差异对比上。通常采用的测试模型如图3所示，首先，从样本库中选取测试测试样本，在标准样纸上打印出来。然后将该测试样本进入被测设备进行识别，同时记录识别结果和响应时间；按要求随机选取若干的测试样本进行测试，并且都记录识别结果和响应时间。

最后根据记录的识别结果和响应时间来计算相应指标。

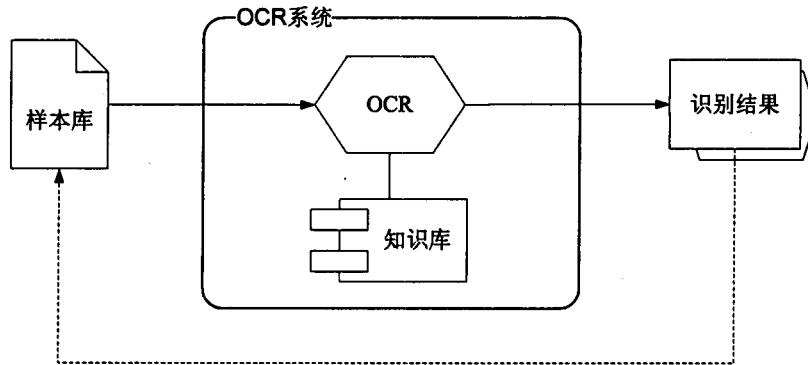


图3 OCR 技术测试模型

7.3.4 语音放大技术评测方法

如无特殊说明，语音放大技术的评测均应在设备额定工作状态下进行。在语音放大设备的输入和输出端分别连接与实际阻抗相同的模拟电阻。

7.3.4.1 频率响应差

在发送端输入300~3400Hz的1/3倍频程点的正弦扫频信号，测量并计算各频率点的灵敏度S，频率响应差即为其他频率点相对于1000Hz处灵敏度的相对值。

7.3.4.2 总谐波失真

在发送端输入频率为300Hz、500Hz和1000Hz的正弦信号，分别测量接收到的谐波信号和基波信号，总谐波失真即为谐波和基波信号的有效值之比。

如音量可调，则应在各个音量范围内测量并计算总谐波失真。

7.3.4.3 信噪比

在发送端输入一定幅度的1000Hz的正弦信号，当总谐波失真≤2dB时，比较此时接收到的信号功率与噪声功率，它们的比值就是以分贝表示的信噪比。



附 录 A  
(资料性附录)  
增强键盘的扩展功能

增强键盘的扩展功能见表A.1。

表 A.1 增强键盘的扩展功能

特 性	提供的功能
粘滞键	提供模式键（如Shift、Control）的观看或闭锁功能，可以避免同时使用按键，用单指按键来实现复合键的操作。即允许用户按下一个像Shift、Ctrl、Alt或Windows Logo（Windows徽标）这样的模式键，并在按下另一个键之前保持其处于激活状态
鼠标键	一种可以供选择的鼠标，用键盘控制实现光标移动及鼠标按钮功能
触控板（Touch Pad）	是一种触摸敏感的指示设备，它可以实现一般鼠标的的所有功能。通过手指在Touch Pad上的移动，能够容易地完成光标的移动。通过按动Touch Pad下方的按键，可完成相应的点击动作（按动左、右键即相当点击鼠标的左、右键）
重复键	延缓重复按键操作的生效时间。使协调时间有限的用户有足够的时间完成按键的释放。因为如果不能快速的抬起手指关闭键盘，会产生意外的字符重复。重复键允许调整重复速率或完全禁用它
慢键	确保一个按键在按键作用生效之前是被持续按下一段时间的。这样可以避免协调性受限的用户意外的按键操作
回弹键	延缓两次有效击键间的时间间隔，确保在接收下一个按键操作前有足够的时间。这样可以避免那些有颤抖症状的用户的无意按键操作
套索键	ToggleKeys功能是专为那些视力受到损伤或存在感知能力残疾的用户而设计的。当开启ToggleKeys选项后，计算机在按下功能锁定键（CAPS LOCK，NUM LOCK或SCROLL LOCK）时发出声音提示。当相应功能键被开启时，系统将发出一个音调较高的提示音；当功能键被关闭时，则发出一个音调较低的提示音

## 参 考 文 献

1. GB/T 21023-2007 《中文语音识别系统通用技术规范》
  2. ITU-T 建议 P.851 基于对话系统业务的主观评测 (ITU-T Recommendation P.851 Subjective Evaluation of Services Based on Spoken Dialogue Systems)
  3. ETR029 为伤残人士提高和适应电信终端和业务提供的接入便利性建议 (ETR029 Access to telecommunications for people with special needs Recommendations for improving and adapting telecommunication terminals and services for people with impairments)
  4. ETSI EG 202 116 ICT 产品和服务的指南：设计为所有人 (ETSI EG 202 116 Guidelines for ICT products and services : Design for All (V.1.2.1))
  5. ETSI EG 202 191 多模式交互，通信和导航指南 (ETSI EG 202 191 Multimodal interaction, communication and navigation guidelines V1.1.1)
-

中华人民共和国  
通信行业标准  
信息终端设备信息无障碍辅助技术的要求和评测方法  
YD/T 1890-2009

\*

人民邮电出版社出版发行  
北京市崇文区夕照寺街14号A座  
邮政编码：100061  
北京新瑞铭印刷有限公司印刷  
版权所有 不得翻印

\*

开本：880×1230 1/16                      2009年8月第1版  
印张：2                                      2009年8月北京第1次印刷  
字数：48千字

ISBN 978 - 7 - 115 - 1845/09 - 87

定价：20元

本书如有印装质量问题，请与本社联系 电话：(010)67114922