

第 1 章 文本信息采编实验

1.1 相关背景知识

1.1.1 文本信息

文本信息是多媒体信息中最常见的一种媒体形式，主要包括数字、字母、符号、控制字符等基本元素。除控制字符外，其他元素均具有内部编码和外部图像两个属性，其中内部编码由 ASCII 码表所定义，每个元素的内部编码在 ASCII 码表（如表 1-1）中是一个 8 位（最高位为 0）的二进制编码；外部图像由点阵库来定义，每个元素的外部图像在点阵库中是一个 $M \times N$ 的点阵。

表 1-1 ASCII 码表

L H	0000	0001	0010	0011	0100	0101	0110	0111
0000	NUL	DLE	SP	0	@	P	`	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	"	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	\$	4	D	T	d	t
0101	ENQ	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	'	7	G	W	g	w
1000	BS	CAN	(	8	H	X	h	x
1001	HT	EM	)	9	I	Y	i	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	K	[	k	{
1100	FF	FS	,	<	L	\	l	
1101	CR	GS	—	=	M	]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	_	o	DEL

当文本信息被输入到多媒体计算机中，就会产生对应的 ASCII 码，并保存在存储器中；当需要输出（显示或打印等）文本信息时，多媒体计算机就从存储器中取出对应的 ASCII 码，

并用此码访问点阵库,读出对应的外部图像显示或打印出来,这样人们就可以看到多媒体计算机内部的文本信息了。在此过程中,文本信息的输出需要有相应的格式才能符合人们的阅读习惯,控制文本信息的输出格式以及传输通信是控制字符的基本作用。

字母与符号按照相应语法规则组合使用,可表示不同的语言文字,文字是文本信息的主体。在我国,多媒体计算机中的文本信息主要分为西文(英文)信息和中文信息,中文信息中的汉字由汉字信息交换码所定义,其中规定一个汉字的编码用两个字节表示,称此编码为汉字内码。

需要特别说明的是,我国少数民族文字(如藏文、蒙文、维吾尔文等)的计算机处理需要专门的信息编码和相应的信息处理系统。本教材不涉及少数民族文字信息处理的内容。

1.1.2 文本信息采集

文本信息的采集是指通过多种途径,快速有效地将不同形式的文本信息输入到多媒体计算机系统中。多媒体计算机中的文本信息采集与一般的文字输入的最主要区别就在于通过“多种途径”,这些途径是由多媒体计算机系统的多媒体处理能力来提供的。如图像的扫描和文字识别能力、声音识别能力等。目前,多媒体计算机中常用的文本信息采集方法主要有键盘输入、手写输入、语音输入和 OCR 输入等。

1. 键盘输入

键盘输入是一种传统的文本输入方法,但仍是随时可用和主要的输入方法。通过键盘,英文信息可直接输入;中文信息则通过不同的中文输入编码(输入法)来完成。常用的输入法有“智能拼音”输入法、“二笔”输入法、“五笔字型”输入法、“智能 ABC”输入法、“全拼”输入法和“微软拼音”输入法等。这些输入法有的基于“字形”进行编码,有的基于“读音”进行编码,有的则是“形音结合”进行编码。一般来说,使用键盘输入中文信息,需要进行专门的技术培训和长期的练习,至少掌握一种汉字输入方法。因此键盘输入法对使用者的计算机操作技能要求较高。

2. 手写输入

手写输入是广泛应用的一种非常人性化的中、英文输入法,适合于不习惯键盘操作的人群和没有标准英文键盘的场合,常用的智能手机、PDA、Tablet PC 等都配备了手写输入系统。

手写输入系统由手写笔、手写板和手写识别软件三部分组成,使用时只要把手写板与电脑主机正确连接,并安装识别软件,即可像真正在纸上写字一样向电脑输入信息。与键盘输入相比,手写输入的最大特点就是操作简单,只要会写字即可输入文字。

手写输入系统不是一般 MPC 的标准配置,因此,需要购买相应的产品。在微软的 MS Office 2003 版中专门增加了手写输入方式,支持一般的手写板输入。如果没有手写板和手写笔,该输入法还可以用鼠标仿真手写笔来进行手写输入。

3. 语音输入

语音输入是通过 MPC 中的音频处理系统(主要包括声卡和麦克风),采集处理人的语音信息,再经过语音识别处理,将说话内容转换成对应的文字来完成输入的。语音输入的最大特点是只要会说话,就能把信息输入到电脑中,但在具体使用之前要经过一个短时间的语音“适应”训练。

语音训练的目的，主要是通过朗读使语音识别软件学习并习惯使用者的发音方式和发音习惯。语音训练需要朗读一段完整的文本，文本内容是由识别软件自行准备的，训练时间大约需要 15 分钟。朗读训练的词句越多，今后语音识别输入时的正确率就越高。所以，如果希望使用语音输入，就要耐心接受语音训练。

训练完成后，就可以用正常说话的语速朗读要输入的文本内容，计算机就会逐字逐句将其识别成文字。如果发现识别错误，可用其他输入方法修改错误的内容。有关语音识别的详细内容，请参见有关技术资料。

与手写输入类似，语音输入系统也不是 MPC 的标准配置，因此也需要购买相应软件。在 MS Office 2003 中也增加了语音输入功能。只要 MPC 具备语音处理能力并装有话筒，就可以通过 MS Office 2003 来使用语音输入功能了。

4. OCR 输入

OCR（光学字符识别）输入是通过识别扫描图像中的文字来达到快速输入文字的目的。这种输入通常是先将印刷品（书本、资料）上的印刷文字通过图像扫描仪以图像方式输入到 MPC 中，再用 OCR 文字识别软件将图像中的文字识别出来，并转换为文本格式的文件，完成文本信息的快速输入。

使用 OCR 输入之前，首先要安装相应的 OCR 识别软件，具体使用过程一般包含以下三个环节：

（1）扫描：通过扫描仪将文本资料扫描成图片输入到计算机。需要注意的是，扫描线数要在 300 线以上，图片格式为非压缩的 TIFF 格式。

（2）纠偏：由于被扫描的文本内容是手工放置在扫描仪平板上的，所以，很难保证被扫描文本与扫描范围的垂直关系，所以可使用识别软件的纠偏校正功能，将扫描得到的图片校正，以便正确识别。

（3）识别：选择要识别的文字区域，开始识别。识别完成后，将识别结果以文本文件的形式保存。如果识别的某些文字有错，系统以颜色区别可能不正确的文字，并提供对照，用户可对照手工修改。

OCR 输入适合于将已经印刷出版过的印刷文字重新输入到计算机中，能够在短时间内输入大量信息，可应用于档案、资料整理和多媒体应用系统的文本输入。

1.2 文本信息的键盘输入实验

1.2.1 实验目的

- 了解多媒体计算机中文本信息输入途径。
- 了解键盘输入文本信息的基本环境组成。
- 掌握键盘输入中、英文信息的基本方法。
- 了解文本信息编辑的基本方法。

1.2.2 实验内容

在 MS Word 软件控制下，通过键盘输入一段中文信息，保存于相应的文件中。

1.2.3 实验环境

- (1) 带有标准键盘的个人计算机。
- (2) 安装支持中、英文处理功能的操作系统（Windows XP/2000/2003、Linux 等），本实验默认操作系统为 Windows 2003 中文版。
- (3) 安装了文字处理软件（可选），本实验默认安装的文字处理软件为 MS Office 2003 中的 Word 字处理软件。
- (4) 用于保存所输入的文本信息位置。本实验默认的文本信息存储位置为：“F:\多媒体技术与信息处理实验\多媒体实验素材\实验 1” 文件夹，具体文件名可自己指定。

1.2.4 实验过程

- (1) 启动个人计算机，进入操作系统操作桌面，如图 1-1 所示。

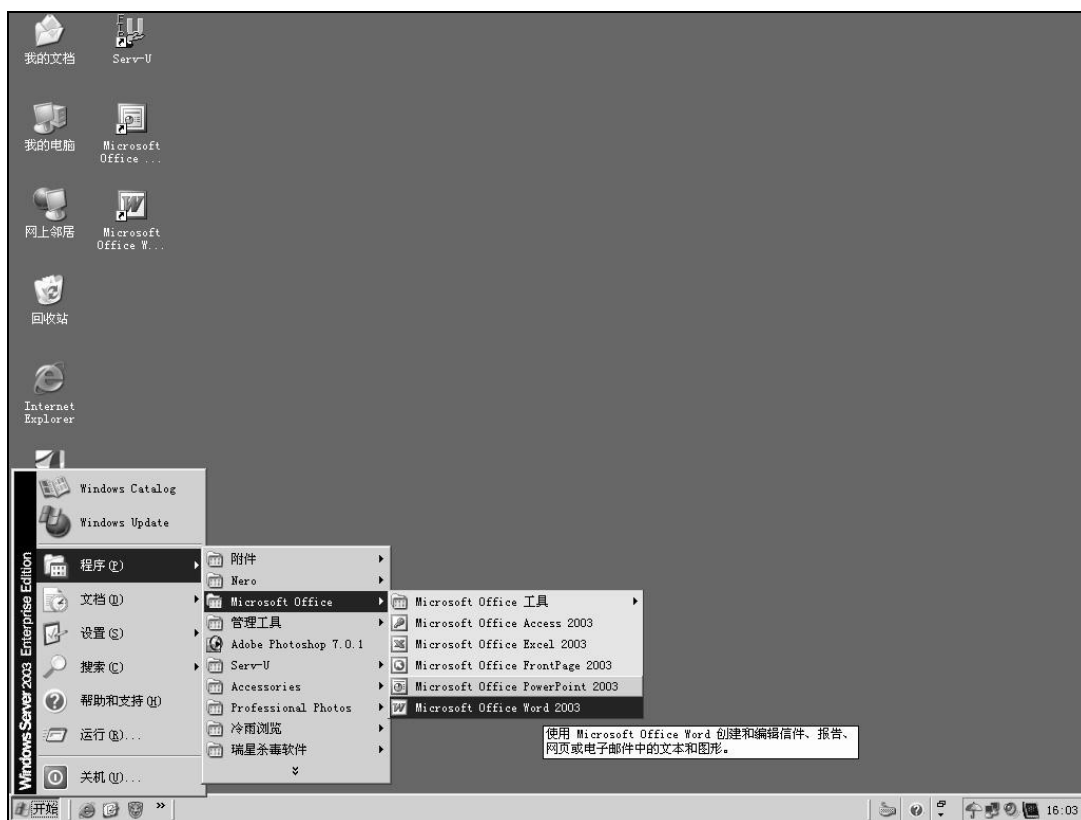


图 1-1 Windows 2003 桌面

- (2) 从“开始→程序→Microsoft Office→Microsoft Office Word 2003”打开文字处理程序，以便于输入和编辑文本信息。MS Word 2003 文字处理程序打开后的屏幕布局如图 1-2 所示。
- (3) 依据教学要求或个人特点，在任务栏选择适当的中文输入法。本例默认选择“微软拼音输入法 2003”。
- (4) 用选定的输入法连续输入以下文字（共 255 个汉字）：

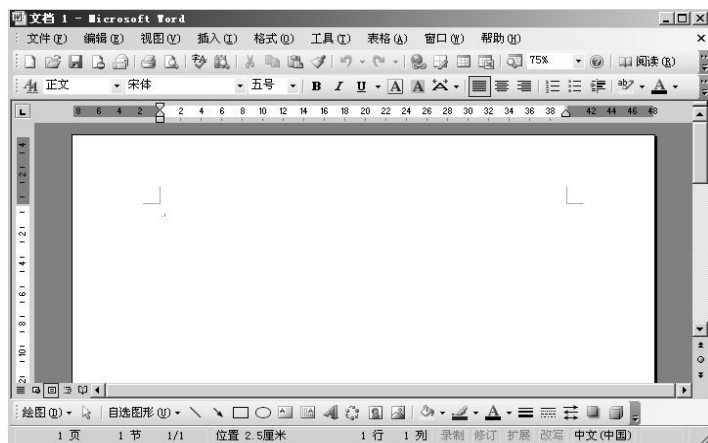


图 1-2 MS Word 2003 文字处理程序的屏幕布局

键盘输入是一种传统的文本输入方法,但仍是随时可用和主要的输入方法。通过键盘,英文信息可直接输入;中文信息则通过不同的中文输入编码(输入法)来完成。常用的输入法有“智能拼音”输入法、“二笔”输入法、“五笔字型”输入法、“智能 ABC”输入法、“全拼”输入法和“微软拼音”输入法等。这些输入法有的基于“字形”进行编码,有的基于“读音”进行编码,有的则是“形音结合”进行编码。一般来说,使用键盘输入中文信息,需要进行专门的技术培训和长期的练习,至少掌握一种汉字输入方法。因此键盘输入法对使用者的计算机操作技能要求较高。

(5) 如图 1-3 所示,将所输入的文字内容保存到本实验环境中的“实验 1.2”文件夹下的 KeyBoardInput.doc 文件中。

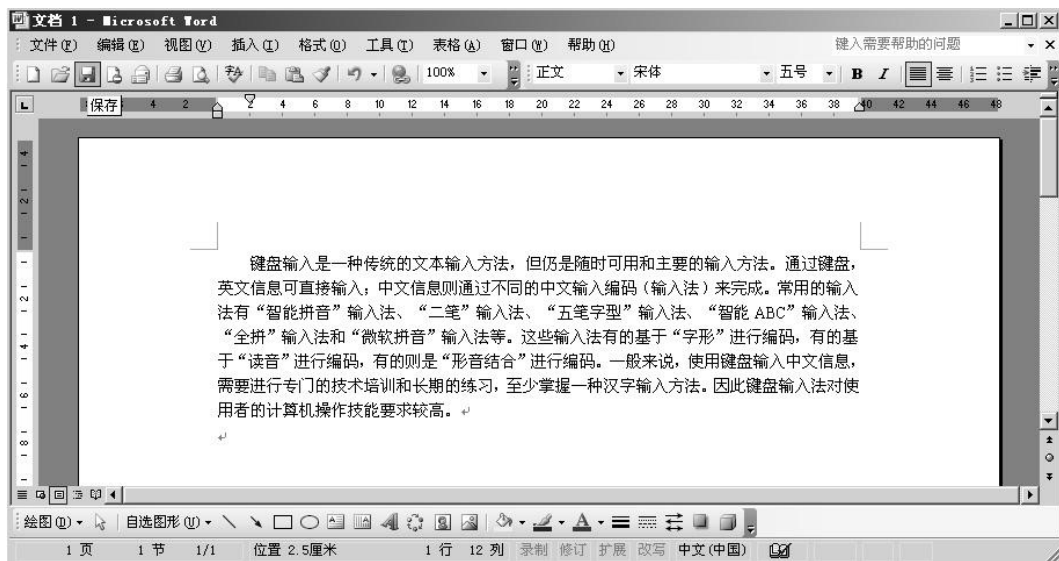


图 1-3 输入了文字的 Word 2003 界面

至此,通过键盘输入文本信息的基本过程已经完成。

1.2.5 实验指导

在以上的实验过程中，有三个主要的技术环节，一是“选择中文输入法”，二是“词汇、短语或语句的快速输入”，三是“选择指定的文件夹”。如果不熟悉某一环节的操作，请参见以下内容。

1. 选择中文输入法

默认情况下，Windows 系列简体中文版操作系统启动后的键盘输入状态为“美式键盘”英文输入，并在任务栏的右侧以黑线框内的图标显示出来，如图 1-4 所示。



图 1-4 任务栏中的“英文输入”状态图标

从英文输入状态进入中文输入状态的操作方法有两种：一种是键盘切换，通过多次按 Ctrl+Shift 键来切换；另一种是通过鼠标操作来完成，用鼠标单击黑线框内最右侧的  图标，可打开中文输入法选择菜单，如图 1-5 所示。

用鼠标指向你所需要的中文输入法并单击鼠标，即可选择相应的中文输入法（如果指向的是“中文（中国）”，则在此进入中文输入法）。此例中，默认选择“微软拼音输入法 2003”。

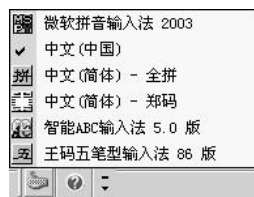


图 1-5 中文输入法选择菜单

一旦进入所选择的中文输入法后，其状态信息将在任务栏的右侧显示出来，如图 1-6 所示的黑线框部分，此时可按照相应的编码方法和要求输入中文信息。



图 1-6 进入中文输入法的任务栏

“微软拼音输入法 2003”的基本编码方案是使用汉语拼音的“全拼”——即每个汉字都要输入其所对应的每个拼音字母，如输入“光”字，需要输入“guang”五个拼音字母。

2. 词汇、短语或语句快速输入

“微软拼音输入法 2003”的最大好处是：支持词汇、短语或语句的快速输入。这种支持能力是通过“智能学习”来实现的。

利用普通的拼音输入法输入中文信息的基本步骤是：先输入拼音，再在许多同音字中挑选自己所需的汉字，然后再输入下一个汉字。而“微软拼音输入法 2003”可连续输入拼音（不用人为选字）由系统根据输入的内容自动选字，从而构成所需的“词汇”、“短语”或“句子”，这大大提高了输入速率。例如：要输入“多媒体计算机技术”，只要连续输入其拼音“duomeitijisuanjijishu”再按空格和回车键即可完成输入。也就是说，利用“微软拼音输入法 2003”完成词汇、短语或语句的快速输入的基本方法是：连续输入拼音后再按空格和回车键。

3. 选择指定文件夹

当第一次保存文件或需要将文件“另存为”其他文件时，都要遇到“选择指定文件夹和文件名”的问题，打开如图 1-7 所示的对话框。只要用鼠标单击所要选择的目标文件夹就可完成指定文件夹的选择。本例中，单击“多媒体实验素材”文件夹，在打开的图 1-8 所示的目标

文件夹中单击“实验 1”文件夹，再单击“实验 1.2”文件夹即选择了最终的文件夹位置。再将文件名更改为“Keyboardinput.doc”，然后单击“保存”按钮即可。

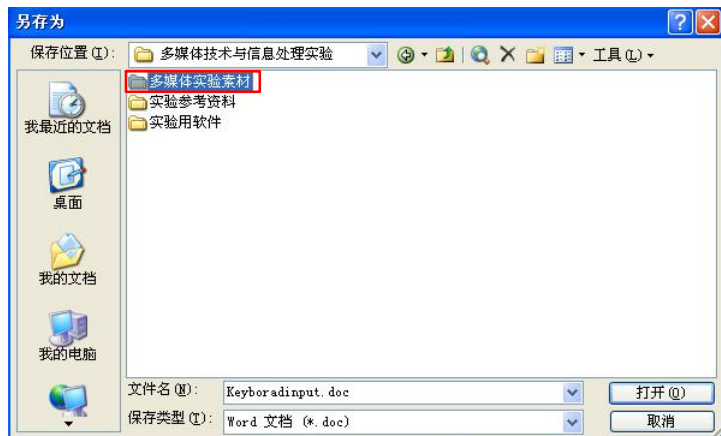


图 1-7 选择文件保存位置对话框（1）

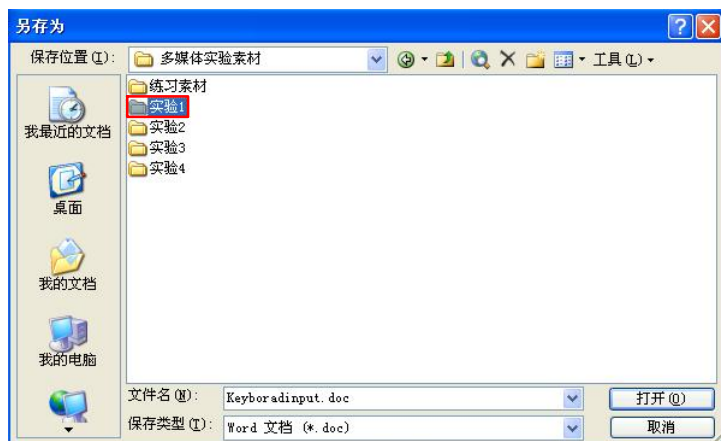


图 1-8 选择文件保存位置对话框（2）

1.2.6 实验结果

本实验的最终结果是：用键盘输入方式在本实验环境的“实验 1.2”文件夹中生成了一个内容为 255 个汉字的名为 Keyboardinput.doc 的文件。

1.3 文本信息的手写输入实验

1.3.1 实验目的

- 了解多媒体计算机中文本信息输入途径。
- 了解手写输入文本信息的基本环境组成。
- 掌握手写输入中、英文信息的基本方法。
- 了解文本信息编辑的基本方法。

1.3.2 实验内容

在 MS Word 软件控制下,通过鼠标模拟手写笔,输入一段中文信息,保存于相应的文件中。

1.3.3 实验环境

(1) 带有标准键盘和鼠标的个人计算机。

(2) 安装支持中、英文处理功能的操作系统 (Windows XP/2000/2003、Linux 等), 本实验默认操作系统为 Windows 2003 中文版。

(3) 安装了 MS Office Professional Edition 2003 组件, 其中包含文字处理软件 Word 2003 和手写输入组件。通常情况下, 手写输入法不是 MS Office 2003 的默认安装选项, 因此要专门安装。

(4) 用于保存所输入的文本信息位置。本实验默认的文本信息存储位置为: “F:\多媒体技术与信息处理实验\多媒体实验素材\实验 1” 文件夹, 具体文件名可自己指定。

1.3.4 实验过程

(1) 启动个人计算机, 进入操作系统操作桌面。

(2) 安装 MS Office 2003 中的手写输入法。具体方法见本实验的实验指导。安装完成后, 系统就具备了手写输入的支持能力, 但任务栏的中文输入状态 (又称输入法工具箱) 并无什么变化, 如图 1-9 (a) 所示。

(3) 单击  图标并选择 “微软拼音输入法 2003”, 使得任务栏中的输入法工具箱变为如图 1-9 (b) 所示的状态。



(a) 英文输入状态



(b) 中文输入状态

图 1-9 英、中文输入法工具箱状态

(4) 单击图 1-9 (b) 中最右边的 “下箭头” 图标, 系统则打开如图 1-10 所示的输入法工具箱内容, 其中多了一个 “输入板”。



图 1-10 输入法工具箱内容

(5) 将鼠标移到 “输入板” 并单击, 此时在中文输入状态中将增加一个输入板图标, 如图 1-11 (a) 所示。当进入相应的文字处理程序后, 任务栏的输入法状态会自动变为带有手写输入图标的状态, 如图 1-11 (b) 所示, 其中黑线框内为手写输入图标。以后每次进入文字处理软件, 系统将会自动打开如图 1-11 (b) 所示的输入法工具箱。



(a) 含有输入板图标的工具箱

(b) 含有手写输入图标的工具箱

图 1-11 有手写输入的输入法工具箱

(6) 从“开始→程序→Microsoft Office→Microsoft Office Word 2003”打开文字处理程序，以便于输入文本信息。MS Word 2003 文字处理程序打开后的屏幕布局如图 1-12 所示。注意右下方的中文输入法工具箱的内容。

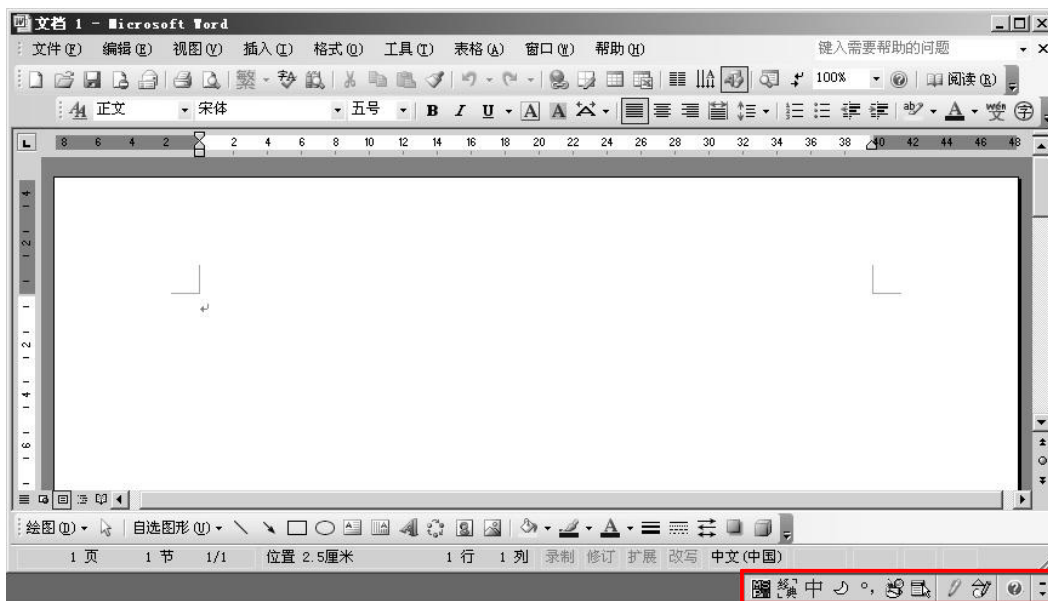


图 1-12 MS Word 2003 文字处理程序的屏幕布局

(7) 在任务栏的输入法工具箱中，单击图标选择“手写输入”，此时系统会打开手写方式选择菜单，如图 1-13 (a) 所示。选择“框式输入”，系统则进入“框式手写”输入状态，如图 1-13 (b) 所示。图中，左右方框均用于手写输入，解决连续手写问题。



(a) 选择框式手写输入

(b) 框式手写输入板

图 1-13 框式手写输入

(8) 用鼠标在左侧方框中书写任意汉字，比如“杨”，经过短暂的识别后在文字处理程序的光标处就会出现所输入的汉字“杨”，紧接着在右侧方框内书写汉字“帆”，如图 1-14 所示，再经过短暂识别后就会在“杨”字后面出现“帆”。当然如果识别有误，可能出现一个相似的其他汉字，如“肠”或其他什么字。

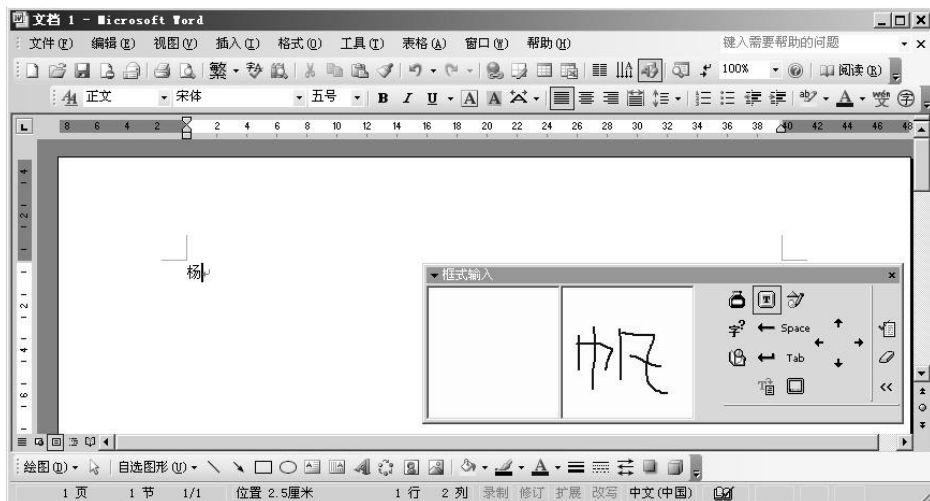


图 1-14 用鼠标手写输入汉字

(9) 为了确保每次输入到文本中的汉字是正确的,可单击“框式输入”板中的查找图标,以便打开相似字浏览查找框。此时,如果在左侧框内写字,右侧框则出现多个相似汉字供选择。如图 1-15 所示,使用者只要从中选出所需的正确汉字即可。



图 1-15 相似汉字的浏览与选择

(10) 连续输入以下文字 (共 80 个汉字)。

手写输入系统由手写笔、手写板和手写识别软件三部分组成。与键盘输入相比,手写输入的最大特点就是操作简单,只要会写字即可输入文字。但手写输入一般要比键盘输入慢很多。

(11) 如图 1-16 所示,将所输入的文字内容保存到本实验环境中的“实验 1.3”文件夹下的 Writing.doc 文件中。

(12) 关闭手写“框式输入”板,退出手写输入方式。至此,通过手写输入文本信息的基本过程已经完成。

1.3.5 实验指导

在以上的实验过程中,有三个重要的技术环节,一是“安装 MS Office 2003 中的手写输入法组件”,二是“中文标点符号的输入”,三是“不同的手写输入方式”。如果不熟悉某一环节的操作,请参见以下内容。

1. 安装 MS Office 2003 中的手写输入法

如果是第一次或重新安装 MS Office 2003,则仅需要在安装过程中选择“自定义安装”,

并在相应的自定义对话框中选择“手写输入”，然后即可开始安装。这一过程比较简单，下面将详细介绍已经安装了 MS Office 2003 但缺少手写输入法的情况下，单独安装手写输入法的基本过程。

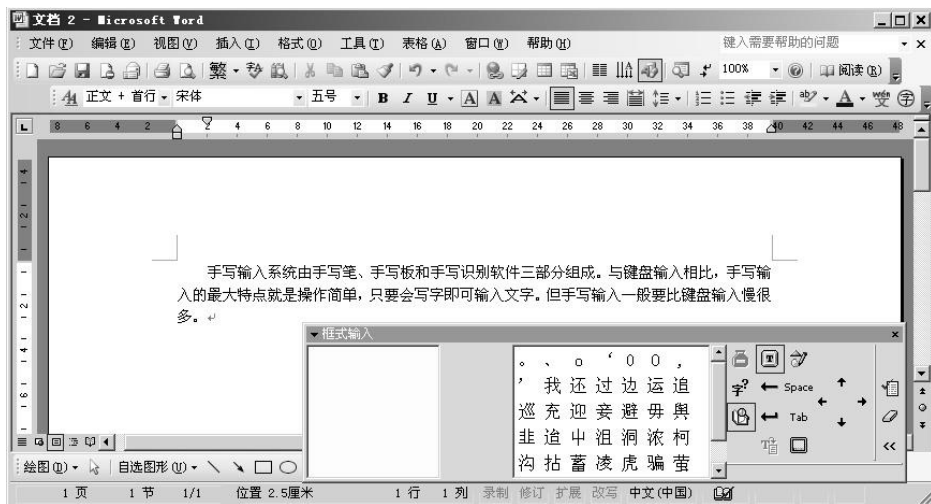


图 1-16 手写输入的文本内容

默认情况下，安装 MS Office 2003 时并不安装手写输入法所需的软件模块，因此，如果所使用的个人计算机系统中仅安装了 MS Office 2003 组件，要想体验手写输入方式而又没有其他专门的手写输入系统时，就需要专门安装 MS Office 2003 中的手写输入法软件。具体安装过程如下：

(1) 启动计算机，进入 Windows 系统（Windows 2003）桌面状态，从“开始→设置→控制面板”打开“控制面板”窗口，如图 1-17 所示。



图 1-17 Windows “控制面板”窗口

(2) 双击“添加或删除程序”，打开“添加或删除程序”窗口，如图 1-18 所示。

(3) 单击“更改或删除程序”，选择 Microsoft Office Professional Edition 2003，再单击“更改”，将打开“Microsoft Office 2003 安装”对话框，如图 1-19 所示。



图 1-18 “添加或删除程序”窗口



图 1-19 “Microsoft Office 2003 安装”对话框

(4) 单击“添加或删除功能”，再单击“下一步”按钮，将打开“Microsoft Office 2003 安装”的“自定义安装”对话框，如图 1-20 所示。

(5) 在“自定义安装”对话框中，选中“选择应用程序的高级自定义”复选框，再单击“下一步”按钮，将打开如图 1-21 所示对话框。

(6) 在“Office 共享功能”旁，单击展开指示器“+”，将打开如图 1-22 所示对话框。

(7) 在“中文可选用户输入方法”旁，单击展开指示器“+”，将打开如图 1-23 所示的对话框。

(8) 在“手写输入”旁，单击下箭头，再单击“从本机运行”，至此，就完成了要安装的“手写输入”软件的自定义安装选项。

(9) 单击“更新”按钮，系统就会按照自定义的内容开始更新系统。

经过以上步骤后，就完成了手写输入法的安装，此时的输入法工具箱就增加了手写输入

工具图标,如图 1-11 所示。



图 1-20 “自定义安装”对话框



图 1-21 Office 组件选择对话框



图 1-22 中文输入法选择对话框

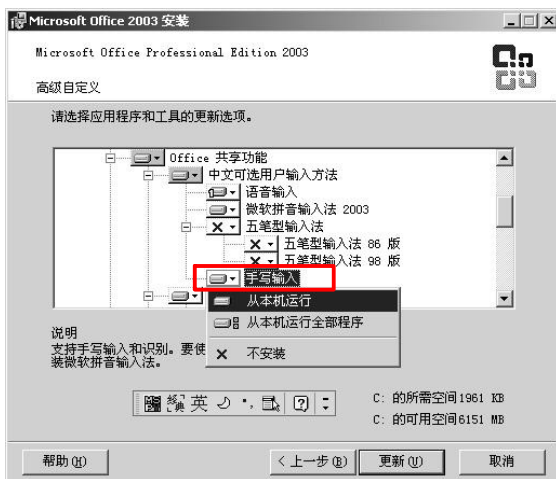


图 1-23 选择安装手写输入法

需要注意的是,在此过程中,系统会提示“需要插入 MS Office 2003 光盘”或“指定 MS Office 2003 安装盘所在的位置”,操作者一定要根据实际情况,找到与目前系统中所安装的版本相同的 MS Office 2003 光盘或磁盘安装文件。否则,安装的手写输入法可能不能正常工作或影响其他输入法的正常使用。

2. 中文标点符号的输入

在 MS Office 2003 的手写输入法中,中文标点如“、。!”等,均可当作汉字一样通过手写来输入。如图 1-16 所示,当输入完一句话或一段文字需要输入句号“。”时,可直接在书写框内书写句号形状,系统会在右侧浏览相似汉字的框内显示可能的标点符号以供选择。使用者只需要选择正确的标点符号(此处为“。”)即可完成标点符号的输入。

另外,英文字母、0~9 十个数字以及其他符号理论上都可以通过手写方式来输入,但由于这些内容笔画少,难以提取准确特征,书写的识别率较低。因此,不建议使用手写输入方式来输入这些内容,而是借助键盘来完成输入比较好。

3. 不同的手写输入方式

事实上, MS Office 2003 的手写输入提供的不仅只有“框式输入”, 还有“任意位置书写”和“绘图板”等手写输入功能, 详见图 1-13 (a)。

所谓“框式输入”, 就是把使用者的书写范围限制在一个指定的方框内, 如图 1-13 (b) 所示。此种输入方式已经介绍, 不再赘述。

所谓“任意位置书写”, 就是不限制使用者的书写位置, 使用者只要进入“任意位置书写”方式, 就可在输入文本窗口的任意位置书写, 经系统识别后的汉字就会插入到当前光标所在位置。具体使用方法如下:

(1) 从图 1-13 (a) 所示的输入方式中选择“任意位置书写”, 此时系统进入“任意位置书写”状态, 如图 1-24 所示。

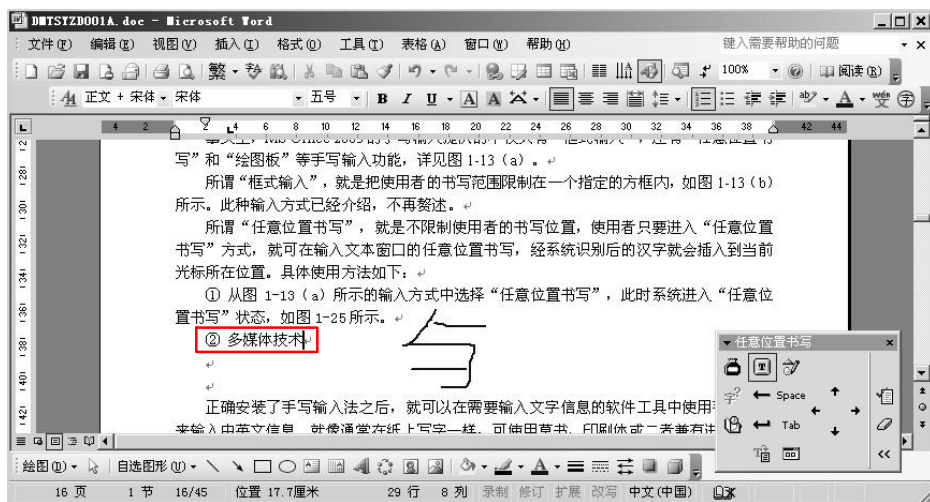


图 1-24 “任意位置书写”输入方式

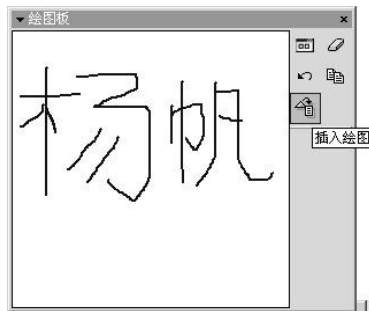
(2) 将当前光标置于希望插入文本内容的位置。

(3) 在文本编辑窗口中的任意位置书写要插入的汉字或符号, 如图 1-24 所示, 此时需要输入汉字“与”。输入完毕, “与”字就被插入到了当前光标所指的“多媒体技术”后面了。

所谓“绘图板”, 就是向使用者提供手绘图能力, 绘出的图画以图像方式插入到光标所在位置, 不需要经过手写识别过程, 如图 1-25 所示, 其中, 图 1-25 (a) 为空绘图板, 图 1-25 (b) 为所绘内容举例。图 1-26 为以图像方式插入的绘图内容。具体使用方法如下:



(a) 空绘图板



(b) 绘图举例

图 1-25 绘图板方式

(1) 从图 1-13 (a) 所示的输入方式中选择“绘图板”，此时系统进入“绘图板”状态，如图 1-25 (a) 所示。

(2) 将当前光标置于希望插入绘图内容的位置，用鼠标在绘图板内绘制任意图画，如自己的签名（此处为“杨帆”），如图 1-25 (b) 所示。

(3) 画好后，单击“绘图板”右侧的“插入绘图”图标，所绘的图画就被插入到当前光标所在位置了，如图 1-26 所示。

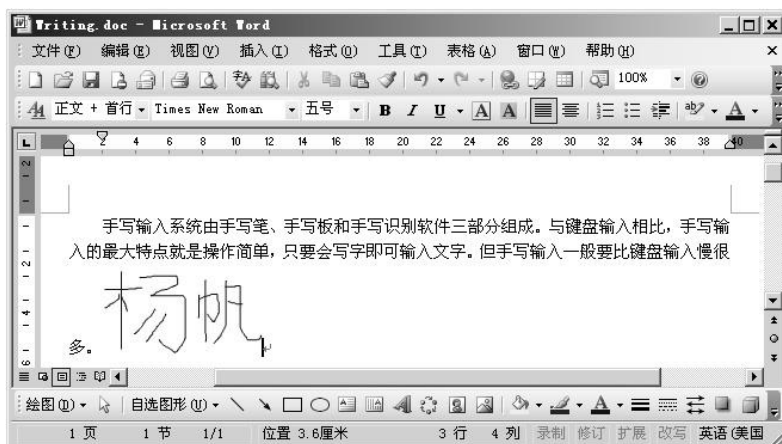


图 1-26 插入了绘图的文本

从以上的使用过程可以看出，“绘图板”功能可用来绘制任意图画、符号、自我签名等，所输入的文字符号是以“手写体”的形式存储和处理的，不需要进行手写识别。而“框式输入”和“任意位置书写”是以“印刷体”的文本形式存储和处理的，需要在书写完成后进行手写识别，从而得到所对应的文字信息。

另外，如图 1-11 (a) 所示，在输入法工具箱中还有一个“输入板”图标，单击该图标会打开一个“输入板—手写识别”窗口，如图 1-27 所示，在“手写输入”方式下，左右两个方框均用作书写区域；在“手写检索”方式下，左侧方框用于手写检索，右侧方框用于显示检索结果，如图 1-28 所示。“输入板”方式与前面所述的“框式输入”很类似，不再赘述。

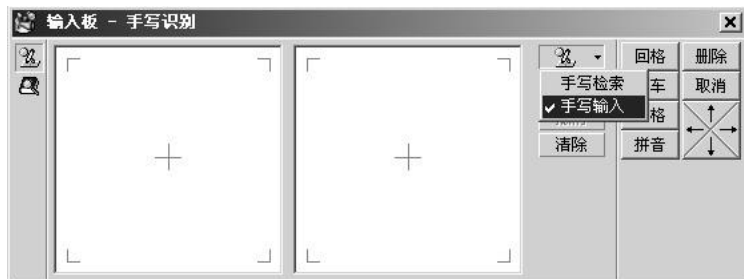


图 1-27 输入板手写输入方式

1.3.6 实验结果

本实验的最终结果是：用手写输入的方式在本实验环境的“实验 1.3”文件夹中生成了一个内容为 80 个汉字的名为 Writing.doc 的文件。



图 1-28 输入板手写检索方式

1.4 文本信息的语音输入实验

1.4.1 实验目的

- 了解多媒体计算机中文本信息输入途径。
- 了解语音输入文本信息的基本环境组成。
- 掌握语音输入中文信息的基本方法。
- 了解文本信息编辑的基本方法。

1.4.2 实验内容

在 MS Word 软件控制下，通过语音识别系统，输入一段中文信息，保存于相应的文件中。

1.4.3 实验环境

- (1) 带有标准键盘、鼠标、声卡、话筒和耳机的多媒体个人计算机。
- (2) 安装支持中、英文处理功能的操作系统（Windows XP/2000/2003、Linux 等），本实验默认操作系统为 Windows 2003 中文版。
- (3) 安装了 MS Office Professional Edition 2003 组件，其中包含文字处理软件 Word 2003 和语音输入组件。通常情况下，语音输入法不是 MS Office 2003 的默认安装选项，因此要专门安装。
- (4) 用于保存所输入的文本信息位置。本实验默认的文本信息存储位置为：“F:\多媒体技术与信息处理实验\多媒体实验素材\实验 1”文件夹，具体文件名可自己指定。

1.4.4 实验过程

- (1) 启动个人计算机，进入操作系统操作桌面。
- (2) 将耳机和话筒插头分别插入多媒体电脑的 SpeakOut 和 Mic 插孔。
- (3) 测试耳机和话筒能否正常工作，即耳机能否听到播放的声音，电脑能否通过话筒录音。具体方法见本实验的实验指导。
- (4) 安装 MS Office 2003 中的语音输入法。具体方法见本实验的实验指导。安装完成后，系统就具备了语音输入的支持能力，任务栏的中文输入状态栏自动出现麦克风图标，如图 1-29 所示。

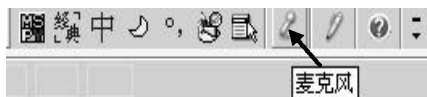


图 1-29 安装了语音输入法的工具箱状态

(5) 单击麦克风图标, 系统会显示如图 1-30 所示的对话框, 提示使用者: 第一次使用语音输入时, 首先要训练语音识别系统, 让语音识别系统熟悉说话者的“口音”, 以便提高识别率。单击“下一步”按钮即可开始训练, 训练过程大约持续 15 分钟。详见本实验的实验指导。

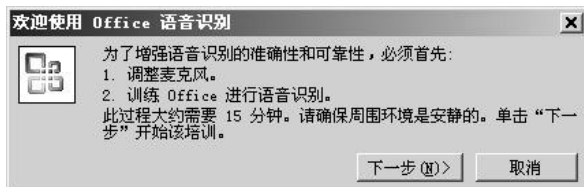


图 1-30 语音识别训练提示对话框

(6) 训练完成后, 系统将进入“听写模式”, 如图 1-31 所示。听写模式右侧为“语音工具”图标, 如果未出现, 可单击最右侧的“下箭头”打开工具菜单, 将语音工具添加进来。



图 1-31 语音输入的“听写模式”状态

(7) 在听写模式下, 只要将光标置于要输入文字信息的位置, 并对着麦克风用训练时的正常语速朗读要输入的内容, 对应的文字就会在光标处出现。比如, 如果对着麦克风朗读“语音识别技术是多媒体计算机技术中的重要技术内容”, 输入的文本行可能出现如图 1-32 所示的输入结果。显然, 图中黑框内的文字应该是“重要”而不是“中医药”。

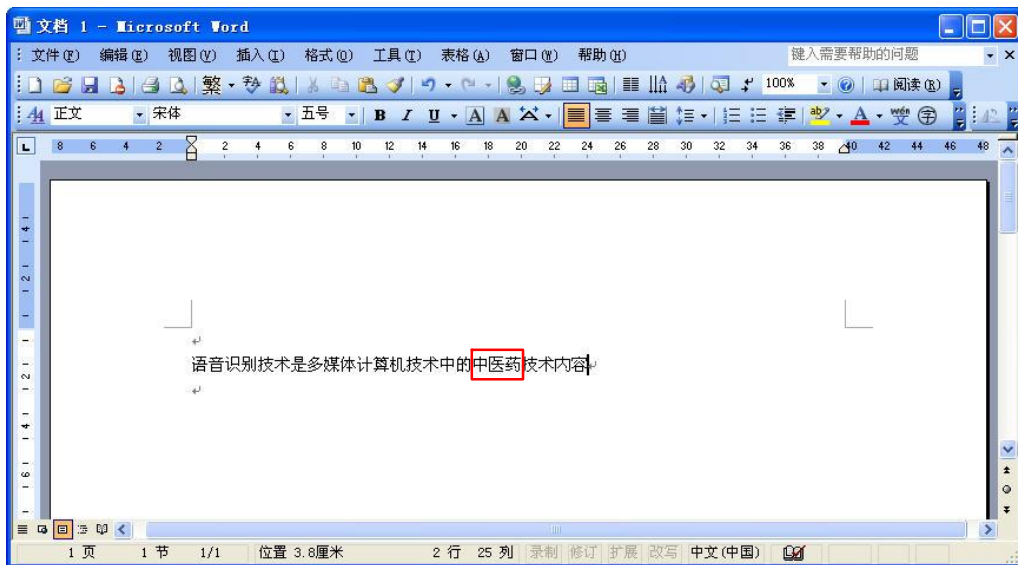
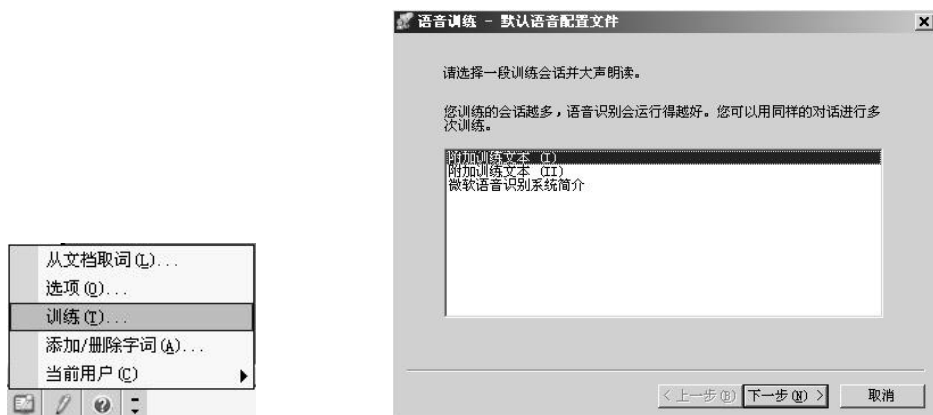


图 1-32 语音输入的文本内容

(8) 为提高识别准确率, 使用者可单击图 1-33 (a) 中的“语音工具”图标并选择“训练”菜单项(在此进行语音适应性训练), 系统将打开如图 1-33 (b) 所示的语音训练内容窗口。使用者可自行选择训练内容进行选练。训练完毕再正式开始使用语音进行中文输入。



(a) 语音工具中的“训练”菜单

(b) 训练内容

图 1-33 训练菜单项及训练内容窗口

(9) 在“听写模式”下用正常语速朗读以下内容(共 103 个汉字)。

语音识别训练是必须的。训练完成后, 就可以用正常说话的语速朗读要输入的文本内容, 计算机就会逐字逐句将其识别成文字。如果发现识别错误, 可用其他输入方法修改错误的内容。有关语音识别的详细内容, 请参见有关技术资料。

(10) 输入完毕, 再次单击“麦克风”图标以关闭“听写模式”, 然后对输入的文字内容进行编辑修改, 正确后保存到本实验环境中的“实验 1.4”文件夹下的 Speech.doc 文件中, 如图 1-34 所示。

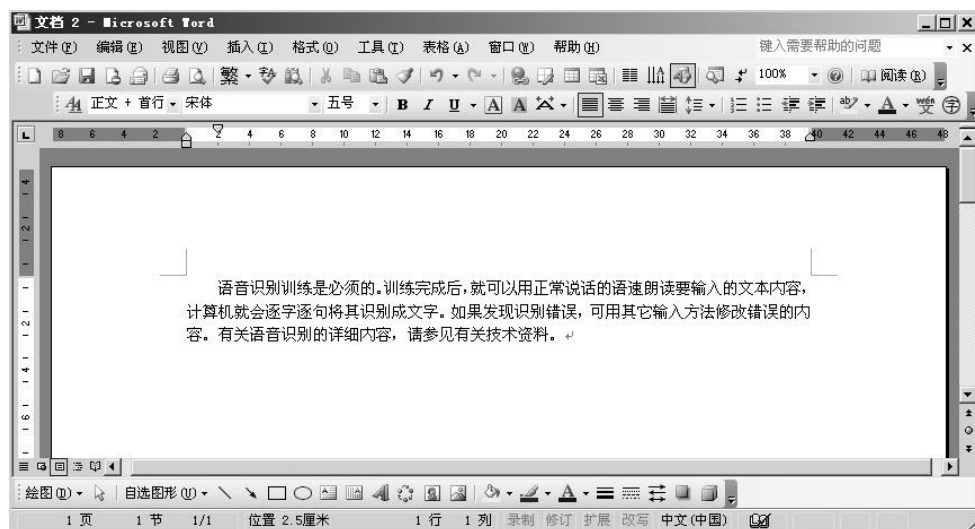


图 1-34 语音输入的文本内容

至此，通过语音输入文本信息的基本过程已经完成。

1.4.5 实验指导

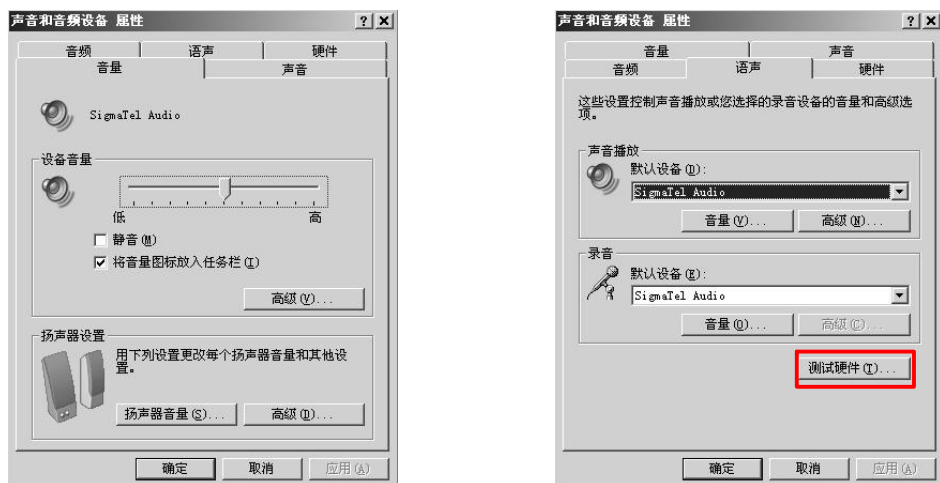
在以上的实验过程中，有四个重要的技术环节，一是“检查测试麦克风和耳机能否正常工作”，二是“安装 MS Office 2003 中的语音输入法组件”，三是“语音识别训练过程”，四是“其他语音功能”。如果不熟悉某一环节的操作，请参见以下内容。

1. 检查测试麦克风和耳机

在计划使用语音识别系统之前，首先要确定自己的多媒体计算机的声音处理部件（声卡及其驱动程序）是否正确安装，这可通过检测麦克风和耳机的工作来完成。具体方法如下：

（1）单击“开始→设置→控制面板”，双击“声音和音频设备”图标，此时，系统会打开“声音和音频设备属性”对话框，如图 1-35（a）所示。

（2）单击“语音”选项卡，系统将转换到如图 1-35（b）所示的对话框。



（a）“音量”选项卡

（b）“语音”选项卡

图 1-35 “声音和音频设备属性”对话框

（3）单击图 1-35（b）中的“测试硬件”按钮，系统将分步骤进行相应的硬件测试，并显示如图 1-36 所示的对话框。



图 1-36 “声音硬件测试向导”对话框

(4) 单击“下一步”按钮，系统首先开始检测与声音信号处理相关的硬件，比如声卡及其接口等，测试进程用渐进的“进度条”表示，如图 1-37 所示，图中的“下一步”按钮在检测过程中是“无效”的，不能使用。如果测试完毕未发现问题，则图 1-37 中的“下一步”按钮变为有效，使用者可继续下面的测试内容。



图 1-37 测试声音硬件对话框

(5) 单击“下一步”按钮，系统将开始进行麦克风测试（也就是录音音量测试）。测试内容主要是确定麦克风是否存在和麦克风的合适音量，如图 1-38 所示。在此过程中，使用者要对着麦克风讲话，比如，朗读图 1-38 中左侧的文字。

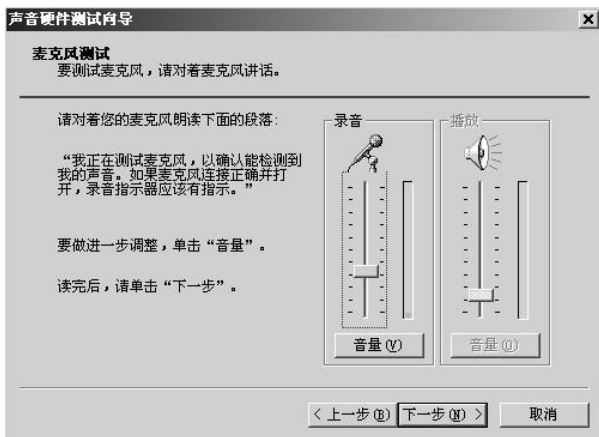


图 1-38 测试麦克风、耳机音量对话框

(6) 测试完毕，单击“下一步”按钮，系统将开始进行耳机测试（也就是播放音量测试），测试内容主要是确定耳机的存在和耳机的合适音量，如图 1-38 的右侧所示，此时右侧的“播放音量”变为有效。如果扬声器或耳机连接正确并打开，播放指示器应该有指示，且能听到自己的声音。在此过程中，使用者也要对着麦克风讲话，比如，“我正在听我自己的声音，如果扬声器或耳机连接正确并打开，播放指示器应该有指示，我应该能听到我自己的声音。”等，这样，可通过耳机听到自己的朗读声。

(7) 测试完毕，单击“下一步”按钮，系统将显示如图 1-39 所示的对话框，表示测试完成。



图 1-39 声音硬件测试完毕对话框

(8) 单击“完成”按钮，退出声音硬件的测试过程，回到系统。

2. 安装 MS Office 2003 中的语音输入法组件

MS Office 2003 中的语音输入法组件的安装与 MS Office 2003 中的手写输入法的安装过程基本相似，其中前(1)至(7)步完全相同，参见 1.3.5 节。

(1) 在图 1-23 中的“语音输入”旁，单击下箭头，再单击“从本机运行”，至此，就完成了要安装的“语音输入”软件的自定义安装选项。

(2) 单击“更新”按钮，系统就会按照自定义的内容开始更新系统。

经过以上步骤后，就完成了语音输入法的安装，此时的输入法工具箱就增加了语音输入的麦克风图标，如图 1-29 所示。

需要注意的是：与安装手写输入法的情况类似，系统也要提示“需要插入 MS Office 2003 光盘”或“指定 MS Office 2003 安装盘所在的位置”，操作者一定要根据实际情况，找到与目前系统中所安装的版本相同的 MS Office 2003 光盘或磁盘安装文件。否则，安装的手写输入法可能不能正常工作或影响其他输入法的正常使用。

3. 语音识别训练过程

刚安装完 MS Office 2003 中的语音输入法软件，必须进行语音识别训练，否则将无法进入“听写模式”。只有充分的语音识别训练，语音系统才会适应使用者的“口音”和发音习惯，这样可以取得较高的识别率。

语音识别训练分为初次训练和附加训练两部分。

初次训练将在系统向导下分步完成，整个训练过程大约需要 15 分钟，具体步骤如下：

(1) 单击图 1-29 中的麦克风  按钮，系统将打开如图 1-40 所示的“麦克风音量调节提示”对话框。

(2) 单击“下一步”按钮，系统将打开语音训练提示对话框，提示使用者在“语句出现时，请按自然、平稳的语调阅读”以及相应的训练方法和技巧，如图 1-41 所示。

(3) 单击“下一步”按钮开始训练，系统显示如图 1-42 所示的对话框，其中包含需要使用者阅读的短句。此时，使用者可平稳读出文本框中的语句，要注意标点停顿。如果遇到不好读的词汇或句子，可单击“跳过单词”按钮而不读；如果希望暂停训练，可单击“暂停”按钮。系统会随着使用者的阅读过程进行识别训练，并用蓝底跟随显示。当读完文本框中的语句后，“下一步”按钮变为有效。



图 1-40 麦克风音量调节提示对话框



图 1-41 语音训练提示对话框

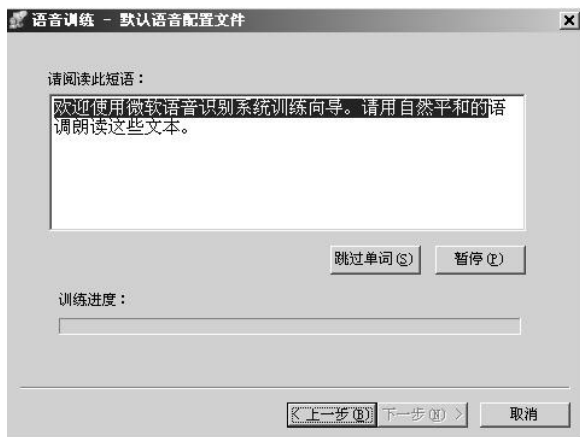


图 1-42 语音训练对话框

(4) 单击“下一步”按钮, 系统则在文本框中显示要训练朗读的下一句话。重复此过程到最后一句要朗读的语句读完。

(5) 单击“下一步”按钮, 系统将打开如图 1-43 所示的对话框, 提示用以上训练的数据

更新语音配置文件。



图 1-43 更新语音训练数据对话框

(6) 更新完毕后，如果对训练结果不满意，还可单击“上一步”按钮继续再训练；如果不再训练，可单击“下一步”按钮结束语音训练过程。

附加语音训练是为了进一步提高语音识别率而设置的。如果在语音输入过程中感觉识别率不高，则可进入附加语音训练过程，具体方法是：单击图 1-33

(a) 中的“语音工具”图标，打开语音工具菜单，选择“训练”即可再次进入语音训练过程，如图 1-33 (b) 所示，其训练内容与初次训练内容不同，使用者可以自行选择。

4. 其他语音功能

当安装了 MS Office 2003 的语音输入法软件之后，输入法工具箱中不仅有语音输入法工具图标，还隐藏有“讲述文本”和“暂停讲述”图标，这两个功能可控制计算机从光标位置开始，以声音形式“念”文本信息，这样使用者就可以通过耳机听到所“念”的文本内容。由于这两个功能图标是隐藏的，所以使用时要先将它们添加到工具箱中。基本过程如下：

(1) 将“讲述文本”图标添加到工具箱。具体方法是单击图 1-31 中最右侧的“下箭头”图标，打开如图 1-44 所示的工具菜单。

(2) 选择“讲述文本”，此时，输入法工具箱中就多了一个“讲述文本”图标，如图 1-45 所示。

(3) 将光标置于要“念”的文本位置，单击“讲述文本”图标，系统将尽可能按句子“念”出这些文字，并将工具箱中的“讲述文本”图标变为“暂停讲述”图标。



图 1-44 工具菜单



图 1-45 带有“讲述文本”图标的输入法工具箱

(4) 使用者带好耳机，听计算机以声音形式“念”出来的文本信息。

(5) 当不需计算机再“念”的时候，可单击“暂停讲述”图标。此时，工具箱中的“暂

停讲述”图标又变为“讲述文本”图标。

1.4.6 实验结果

本实验的最终结果是：用语音输入的方式在本实验环境的“实验 1.4”文件夹中生成了一个内容为 103 个汉字的名为 Speech.doc 的文件。

1.5 文本信息的 OCR 输入实验

1.5.1 实验目的

- 了解多媒体计算机中文本信息输入途径。
- 了解 OCR 输入文本信息的基本环境组成。
- 掌握 OCR 输入中文信息的基本方法。
- 了解文本信息编辑的基本方法。

1.5.2 实验内容

- (1) 扫描一页文字页。
- (2) 用 OCR 软件识别出文字。
- (3) 将文字信息复制到 MS Word 中。
- (4) 保存到相应的文件中。

1.5.3 实验环境

- (1) 带有标准键盘、鼠标的多媒体个人计算机，安装支持中、英文处理功能的操作系统（Windows XP/2000/2003、Linux 等），本实验默认操作系统为 Windows 2003 中文版。
- (2) USB 接口的办公用扫描仪，至少 A4 幅面。
- (3) USB 信号线，配套的扫描仪驱动程序。
- (4) 专门的 OCR 识别软件，如“方正 OCR”、“尚书 OCR”等。本实验使用方正 OCR 软件。
- (5) 选装了 MS Office 2003 组件中的文字处理软件 Word 2003。
- (6) 要输入（被扫描识别）的文稿内容，如一篇文章或一本书等。
- (7) 用于保存所输入的文本信息位置。本实验默认的文本信息存储位置为：“F:\多媒体技术与信息处理实验\多媒体实验素材\实验 1”文件夹，具体文件名可自己指定。

1.5.4 实验原理

1. 扫描系统

扫描系统由扫描仪、多媒体计算机、特定的连接线、驱动软件和扫描软件、图像处理软件、文字识别软件以及图像输出设备（如打印机等）组成，用来提供图像扫描、处理、文字识别、存储、输出等数字化处理能力。

2. 扫描仪的工作原理

扫描仪是一种图像输入设备，可将纸质印刷品图像（如照片、文字页、图形和插画等）

或实物图像（如硬币或纺织品等三维物件）转换为计算机可以处理（显示、编辑、存储和输出）的数字格式图像。其工作原理是：对原稿进行光学扫描，然后将光学图像传送到光电转换器中变为模拟电信号，再将模拟电信号变换成为数字电信号，最后通过输出接口送至计算机中。

扫描仪扫描图像的步骤是：首先将要扫描的原稿正面朝下铺在扫描仪的玻璃板上，原稿可以是文字稿件或者图纸照片；然后启动扫描仪驱动程序后，安装在扫描仪内部的可移动光源开始扫描原稿。为了均匀照亮稿件，扫描仪光源为长条形，并沿 y 方向扫过整个原稿；照射到原稿上的光线经反射后穿过一个很窄的缝隙，形成沿 x 方向的光带，又经过一组反光镜，由光学透镜聚焦并进入分光镜，经过棱镜和红绿蓝三色滤色镜得到的 RGB 三条彩色光带分别照到各自的 CCD 上，CCD 将 RGB 光带转变为模拟电子信号，此信号又被 A/D 变换器转变为数字电子信号。至此，反映原稿图像的光信号转变为计算机能够接受的二进制数字电子信号，最后通过串行或者并行等接口送至计算机。扫描仪每扫一行就得到原稿 x 方向一行的图像信息，随着沿 y 方向的移动，在计算机内部逐步形成原稿的全图。

3. OCR 的工作原理

OCR 是一种典型的扫描系统，它由扫描仪、多媒体计算机、特定的连接线、驱动软件、扫描软件和文字识别软件组成，用来提供文稿扫描、文字识别、校对、存储等快速文本采集能力。OCR 的基本工作原理是：先将文稿以图片形式扫描输入到计算机，再由 OCR 识别软件对选定区域的文字进行识别，得到可编辑的文本信息，如果识别有错，可对照图形文字修改对应的出错文字，最后以可编辑的文本形式存储。

1.5.5 实验过程

1. 安装 OCR 软件

- (1) 用随机带的 USB 线将扫描仪和电脑连接起来，并打开扫描仪电源开关。
- (2) 启动个人计算机，进入操作系统的操作桌面。
- (3) 安装扫描仪驱动程序，具体方法见扫描仪使用手册。
- (4) 安装方正 OCR 识别软件，具体方法如下：

1) 定位到方正 OCR 软件，运行 SETUP.EXE，进入安装过程，屏幕显示图 1-46 所示画面。



图 1-46 方正 OCR 软件安装提示画面

2) 初始设置完成后，进入图 1-47 所示画面。输入产品序列号，单击“下一步”按钮，系

系统将进入如图 1-48 所示的画面。



图 1-47 方正 OCR 序列号输入对话框

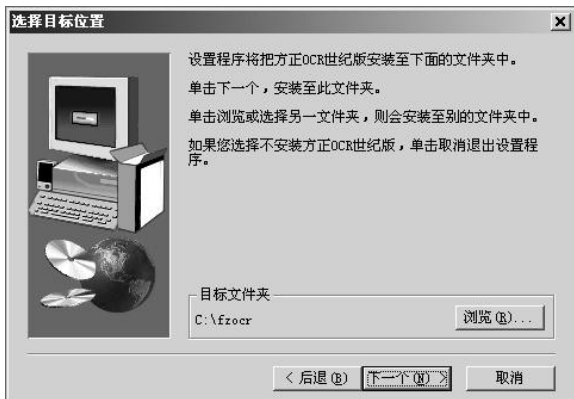


图 1-48 指定软件安装位置对话框

3) 若需指定软件安装位置,可单击“浏览”按钮并选择合适的文件夹;若使用默认位置,则可直接单击“下一步”按钮开始具体的软件安装工作。

4) 安装完毕,系统会提示如图 1-49 所示的对话框。



图 1-49 方正 OCR 安装完成对话框

5) 单击“结束”按钮,完成方正 OCR 的安装过程。此时会在 Windows 桌面和“程序”菜单中出现“方正 OCR 世纪版”图标。

2. 运行方正 OCR

开始文字的扫描识别工作。具体方法是单击“开始→程序→方正 OCR 世纪版”,此时进入方正 OCR 识别程序,如图 1-50 所示。完整的 OCR 输入过程包括扫描、纠偏、切分、识别、校对、保存文件等主要环节。

(1) 扫描。一页内容的扫描过程可分为预扫描、设置参数、选择区域、扫描四个环节。

1) 预扫描:将要识别的印刷品(书稿、讲稿等)放在扫描仪的玻璃板上,调整位置,尽可能使之与扫描仪的参考线呈平行或垂直状态,用左手轻轻按住印刷品以免位置发生移动,右手单击图 1-51 中左下部的“扫描”按钮,此时启动扫描仪,开始预扫描。预扫描完毕,被扫描的页面内容就呈现在扫描区域中,如图 1-51 右侧所示。

2) 设置参数:单击图 1-51 右上部线框内的“图像类型”右侧的下箭头,从下拉列表中选

择“黑白文稿”，再单击“目标设备”右侧的下箭头，从下拉列表中选择“OCR”，此时“分辨率”右侧的参数会自动修改为“300”线，如果不是，请用鼠标单击右侧的下箭头，并选择300。如果以上三个参数设置有误，将直接影响OCR的识别率。

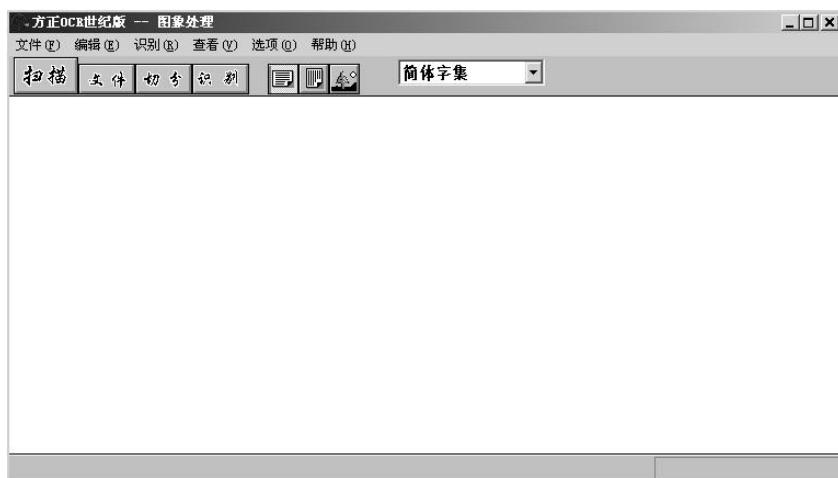


图 1-50 方正 OCR 主画面



图 1-51 扫描程序画面

3) 选择区域：在图 1-51 右侧的扫描区域中，将鼠标从要选择区域的左上角拖至右下角，此时要被扫描的区域就用虚线框起来了。

4) 开始扫描：单击图 1-51 下方的“扫描”按钮，扫描仪则只扫描虚线框内的内容，并显示图 1-52 的进度提示。

扫描结束后，扫描结果将送给 OCR 程序，如图 1-53 所示。

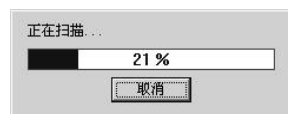


图 1-52 扫描进度提示

为了操作安全,可先将扫描结果保存为.TIFF 格式的图片文件,以后识别处理时则不必再次扫描,直接打开.TIFF 图片就可以进入下面的处理、识别环节。

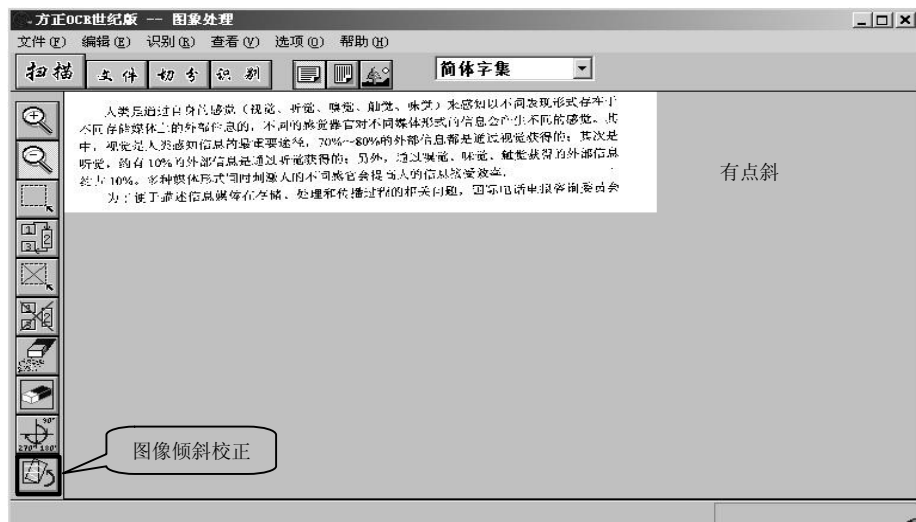


图 1-53 OCR 得到的扫描结果

(2) 纠偏。单击图 1-53 中的“图像倾斜校正工具”将图 1-53 中的文字图片进行角度校正,系统会自动检测需要校正的角度并提示给用户,且允许用户修改(通常不修改),使用系统检测到的校正角度得到的校正结果如图 1-54 所示。

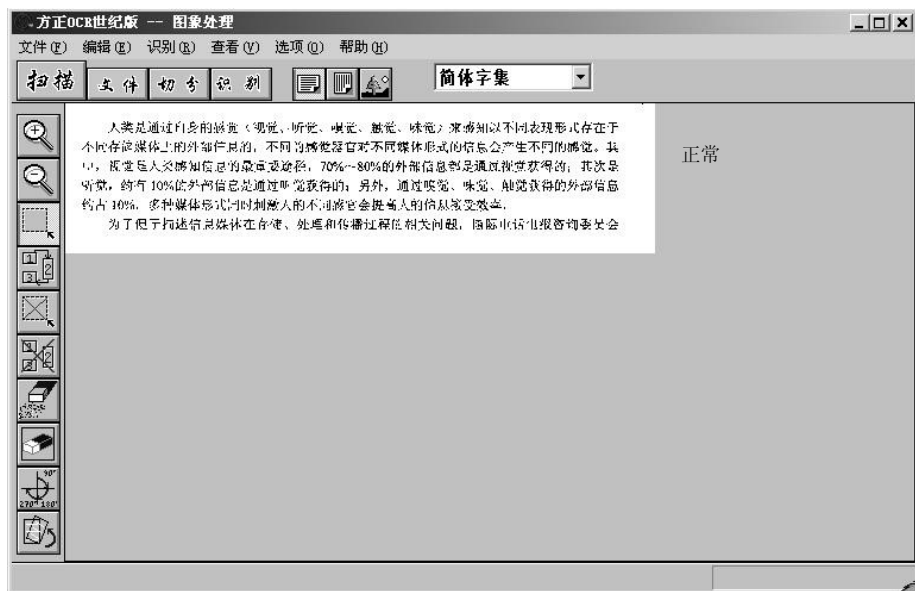


图 1-54 纠偏后的识别对象

(3) 自动切分。单击图 1-54 中的“切分”按钮,系统将对扫描图片按“特征”自动进行区域“切分”,此例中将扫描图片切分为一个区域(或不切分),如图 1-55 中的线框所示。

(4) 识别。单击图 1-55 中的“识别”按钮,系统提示“正在识别”。识别完毕,系统会显示“文稿校对”窗口,允许操作者对识别内容进行校对修改,如图 1-56 所示。

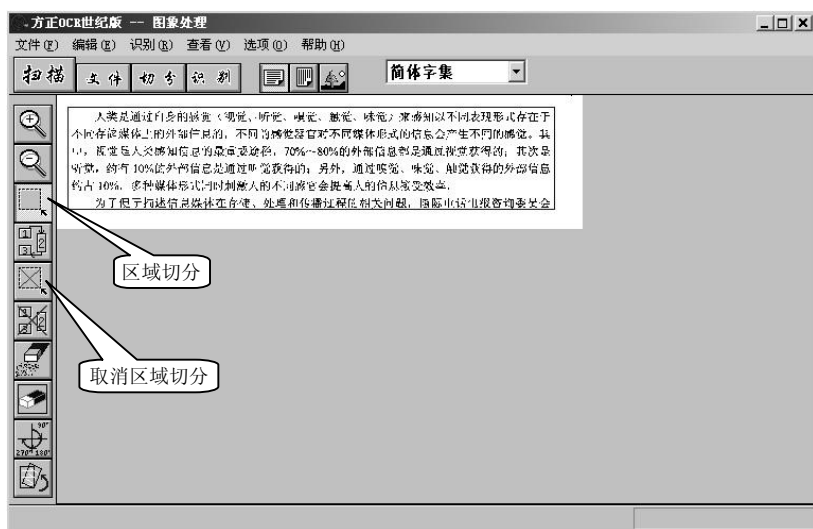


图 1-55 识别对象的区域“切分”结果

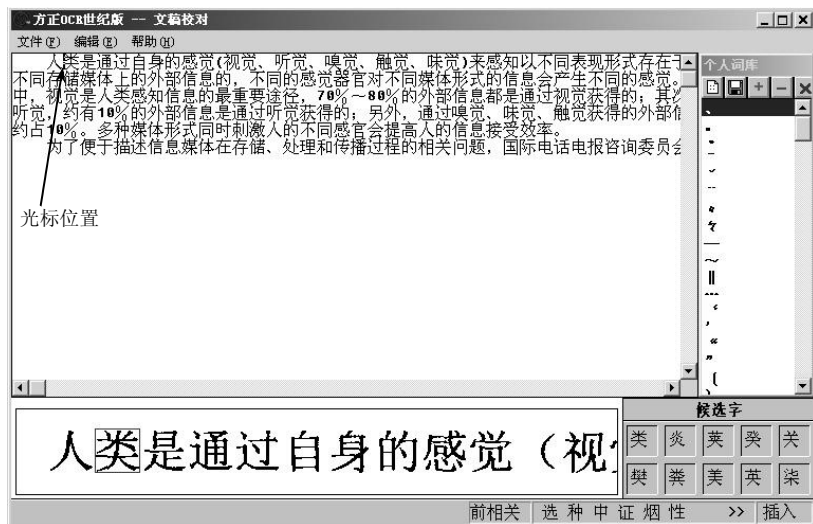


图 1-56 方正 OCR 的“文稿校对”窗口

该图左上方为识别得到的可编辑文字，右上方为“个人词库”，左下方的“大字体”则显示了从“可编辑文字”区当前光标开始的文字内容对应的识别前的“图片文字”，黑框内的图片文字应该是“可编辑文字”区光标所指文字，右下方为可能的“候选字”区。

(5) 校对。在图 1-56 中，将鼠标指向识别有误的文字处，左下方就会显示对应的“图片”文字，人工从“候选字”区选择正确的字或键盘输入正确的字，即可完成对识别有错的文字的校对。

(6) 保存。单击图 1-56 “文件”菜单中的“存储文本文件”或“另存为”命令，即可打开文件“另存为”对话框，此时选择到该实验环境中的“实验 1.5”文件夹下，将文稿内容存储为 OCR-Input.txt 文件。也可将图 1-56 中的文本信息复制到 MS Word 中，进行编辑排版后保存为 OCR-Input.doc 文件。

至此,通过 OCR 识别技术输入文本信息的基本过程已经完成。

1.5.6 实验结果

本实验的最终结果是:通过扫描仪和 OCR 识别技术在本实验环境的“实验 1.5”文件夹中生成了一个内容为 233 个汉字的名为 OCR-Input.txt 或 OCR-Input.doc 的文件。

1.5.7 注意事项

在以上实验过程中,需要重点注意以下几个问题:

(1) 扫描和 OCR 识别是两个既相互联系又相对独立的环节。一方面,扫描质量的优劣直接影响识别率,另一方面,OCR 识别程序中的“角度校正”、“切分”等处理功能可实现对扫描图像的优化处理,从而提高识别率。实际工作中,可采用“批量扫描批量识别”的工作模式,即将要识别的所有文字页批量扫描成图片,并保存为.TIFF 格式的文件。然后,进入 OCR 识别系统,打开.TIFF 图片进行相关处理和识别,再将识别出的文字结果编辑保存为相应的文本文件。这样做的好处是缩短时间,提高速度,进而提高工作效率。

(2) 正确的扫描参数是保证扫描质量的前提。通常情况下,需要用户特意设置的主要是与被扫对象有关的参数,主要包括图 1-51 中黑线框内的“图像类型”、“目标设备”和“分辨率”三个参数。为了满足 OCR 识别的要求,以上三个参数的取值分别为“黑白文稿”、“OCR”和“300”,否则将会影响 OCR 的识别率。

(3) 对于扫描的特殊文稿图片(如分栏目录、分栏期刊、图文表混合文稿等),识别前应先使用“设定识别区域”、“删除识别区域”、“设定识别顺序”、“删除识别顺序”等工具,对识别对象进行人工区域切分和识别顺序设定,这样可有效提高识别率和文稿的可用性。比如,对于“分栏目录”类型的文稿,如果按照图 1-57 所示进行切分和顺序设定,识别结果可直接使用,否则,将“按行”识别导致识别结果无法直接使用。

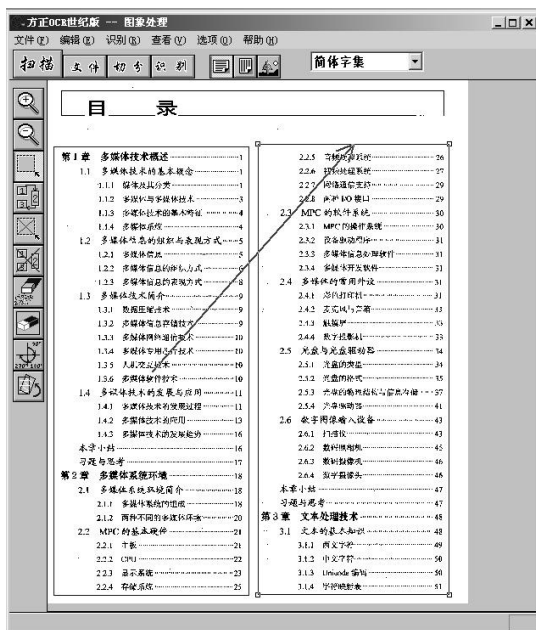


图 1-57 分栏目录的切分与识别顺序

(4) 建议使用好的原稿, 如果原稿文字不清, 有重影或有污损或感光很差, 则很难得到较高的扫描质量和识别率。同时要保持扫描仪的洁净, 扫描之前如果发现扫描仪的玻璃面脏了, 请用干净的软布沾玻璃清洁剂擦拭即可。

1.6 文本编辑实验

1.6.1 实验目的

- 了解文本信息编辑主要内容。
- 掌握 MS Word 2003 文本编辑的基本方法。
- 掌握一般编辑软件的使用方法。

1.6.2 实验内容

- (1) 打开文本文件, 添加、插入、删除文本信息。
- (2) 使用剪贴板(选择、复制、粘贴、剪贴等)。
- (3) 格式编辑(字体、字号、颜色、对齐、分栏等)。

1.6.3 实验环境

- (1) 带有标准键盘、鼠标的多媒体个人计算机。
- (2) 安装支持中、英文处理功能的操作系统(Windows XP/2000/2003、Linux 等), 本实验默认操作系统为 Windows 2003 中文版。
- (3) 安装了 MS Office Professional Edition 2003 组件, 其中包含文字处理软件 Word 2003 和语音输入组件。
- (4) 用于保存所编辑的文本信息位置。本实验默认的文本信息存储位置为: “F:\多媒体技术与信息处理实验\多媒体实验素材实验1”文件夹, 具体文件名可自己指定。

1.6.4 实验过程

- (1) 用 MS Word 2003 打开要编辑的 txt 文件(或进入 MS Word 2003 用键盘输入要编辑的文本)。
- (2) 使用插入、删除功能编辑文本。
- (3) 使用区域选择、复制(剪切)、粘贴功能编辑文本。
- (4) 其他编辑。如字体、字号、颜色、对齐、分栏等。

1.6.5 实验结果

本实验的最终结果是: 通过 MS Word 2003 编辑完成并保存于实验环境的“多媒体实验素材实验1”文件夹中的一个.doc 文件。

1.7 实验小结

本组实验主要包括文本信息的键盘输入、文本信息的手写输入、文本信息的语音输入

以及文本信息的 OCR 输入、文本编辑等五个实验。通过本组实验，学生首先应该了解在多媒体计算机系统中，有多种输入文本信息的途径，这些途径适用于不同的场合与不同的用户。键盘输入是多媒体计算机中文本信息输入的根本途径，手写输入和语音输入是文本信息输入的辅助手段，OCR 输入可用于快速输入大量规范的印刷文字信息。其次，应该掌握至少一种文本信息的输入方法，同时掌握基本的文本信息编辑、排版技术，为今后使用计算机系统奠定基础。