

# 基于语音识别的计算机声图系统

薛啸宇 柳玉昌 赵金龙

(电子技术系)

**摘要** 本文提出一种计算机声图系统的模型,阐述了它的原理,并以 TI-SPEECH 语音处理系统、KX-600 图形显示系统和 IBM PC/AT 计算机为基础,实现了一个简单的计算机声图系统。该系统的图形软件采用在 KX-600 图形显示系统上实现的 MGP 二维军用图形包,实现了语音输入的透明键盘支持。本文还提出了用自然语言输入和 GKD-PROGKS 图形系统构造的智能计算机图形系统的功能模型。

**关键词** 语言信号处理, 显示, 微型计算机, 透明键盘、声图系统

**分类号** TP391.42

近年来,由于人们对语音和自然语言的基本性质有了进一步的了解,特别是语音信号、信息处理技术和大规模集成电路的迅速发展,使人与机器之间用口语交换信息,控制机器运行已逐步成为现实。国内外相继推出了多种计算机语音识别和语音合成产品,利用这一成果可以形成种种实用声控系统,如声控计算机、声控打印机系统、声控作图系统等。

如果用语音设备代替传统的键盘、光笔等设备进行图形命令和数据的输入和输出,可使计算机图形系统智能化,形成一种声控作图的计算机声图系统。

根据目前的技术水平,研制一个以有限词汇量的单词、词组语音识别为基础的命令—图形系统,既具有先进性,又具有实用价值。这种计算机声图系统的功能模型如图 1 所示。其中主要模块为语音处理模块和图形核心模块。

计算机声图系统是一门较新的技术,它有着广阔的应用前景,可以在军事指挥系统、CAD 辅助设计、办公自动化、网络系统、三维地形地图系统、动画系统、等领域大显身手。

## 1 计算机声图系统原理

用语音控制计算机作图,实现语音与图形处理器交互作用,要有三个子系统衔接(如图 2):语音识别输入系统,语音合成输出系统,计算机图形处理系统。

这三个子系统在计算机操作系统的统一管理下运行。为了协调地工作,除了要有相应的硬件支持

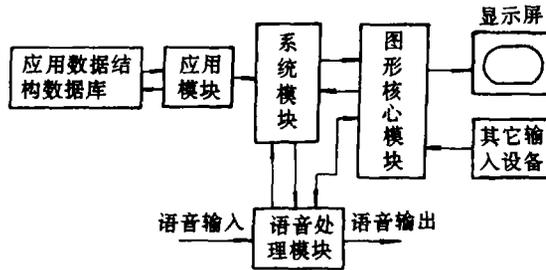


图 1 计算机声图系统模型

\* 1991年4月30日收稿

外，还要有软件实现语音输入到图形的转换，这个转换过程由主机上运行的软件完成。可以认为，该声图系统的各子系统都作为主计算机的外部设备，在主机操作系统支持下工作。它们都通过接口板与主机总线相连，其相应的软件都处于操作系统的上层，因此设计时，不需更多的修改系统软件，从而形成一种层次结构。

语音识别根据不同要求，有多种不同的方法。比较实用的是有限词汇量、固定讲话人的词组语音识别。在用词组语音识别法识别之前，要先对计算机进行训练，训练时抽取各个语音波形参数，建立词组的参考模板，若干个词组的参考模板组合在一起形成一个词库。词库（或词表）的大小取决于应用的需要、识别精度和计算机存储量的大小。识别时将提取到的待识词组的语音参数的词库中的参考模板进行匹配，当误差小于某一设定的门限时，就认为该词组已被识别。这个门限称为拒识门限。

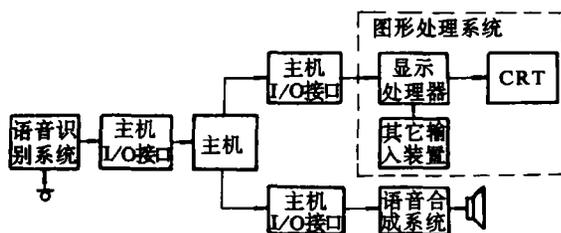


图2 声图系统基本组成

语音合成指用专用的硬设备或计算机再现人能够听得懂的语音波形，国内外在语音合成方面取得的成果显著，已经有声音质量接近于人类自然语言的语音合成系统。

语音合成的方法很多，最简单的是录音编辑方式。但目前研究方向为音素（音节）合成法，这种合成系统可以合成无限词汇量的语音，且占用存储器的容量小，但合成语音质量稍差。

语音合成的方法很多，最简单的是录音编辑方式。但目前研究方向为音素（音节）合成法，这种合成系统可以合成无限词汇量的语音，且占用存储器的容量小，但合成语音质量稍差。

计算机图形系统是在操作系统支持下工作的。它具有图3所示的层次结构。

图形软件包是图形核心系统（GKS）的一种语音界面，它实现图形数据的输入和输出，不同坐标系之间的坐标变换，图形的产生和消去，以及图形的存贮原理等功能。对声图系统还要考虑语音控制特点。因此，应在图形包中扩充语音处理模块，包括语音输入，声图转换、语音输出以及控制等四个部分。

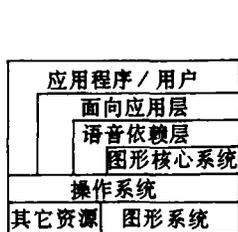


图3 计算机图形系统的层次结构

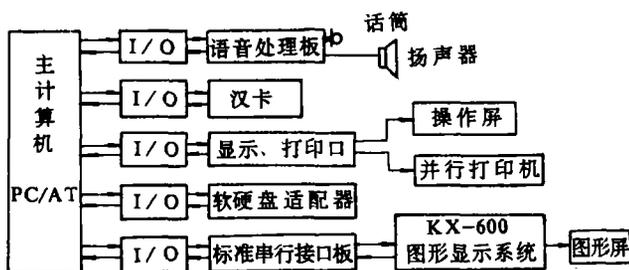


图4 声图实验系统硬件结构

## 2 计算机声图实验系统的设计

为了使系统能适应各种需要，应配置比较齐全的硬件设备，综合考虑系统功能要求、硬件价格、接口实现方法难易程度，调研计算机语音识别和显示设备的内存、速度、容量、处理能力及相互之间的匹配等指标，提出了声图系统硬件结构和软件结构的方案。在此基础上，首先实现较实用的计算机声图实验系统。声图实验系统的硬件结构如图4所示。

TI-SPEEC I 语音处理器是美国德克萨斯仪器公司1986年的产品，它采用80年代最新的数字信号处理研究成果——TM320系列高速数字信号处理器，加上有效的语音识别、合成算法。在IBM-PC系统微型机和TI的专用微机TI-PNSINISE-PRO上实现了一个完整的处理系统。它不仅能够识别词组，也可以进行连续语音识别和词法、句法分析，而且还具有语音编辑录音合成、文字合成（英语）、电话

接口等功能，它是一个功能比较齐全的语音处理系统。

KX-600 图形显示系统是我校研制的硬件。它实际是一个小型工作站，具有四个图形通道，可带四个图形屏。图形分辨率为  $720 \times 512$ ，可前景用图，也可背景作图。甚至前、背景迭加，有 10 种线型，256 种颜色。可进行旋转、放大、平移、开窗、图形的拾取和消去。填充图案可达 174 种。

该显示系统配备了 GKS (2D) 国际标准图形包，MGP 军用图形包，BGP 基本图形包，并配有中英文全屏编辑软件，对话作图软件等。

MGP 是一个二维的军用图形软件包，它适用于作战标图、军事指挥系统。MGP 是按国际 GKS 图形标准在规格化坐标内实现的 C 语言的二维图形库。它具有输入、输出原语及属性控制、坐标变换、段存贮、图形编辑、地图拼接及询问功能，并可实现工作站的控制。整个功能采用 C 语言实现。

### 3 声图实验系统中的关键技术

在形成声图实验系统过程中，KX-600 图形显示系统、图形包的研制 TI-SPEECH 处理系统的使用虽有很多难点，但由于我们多年从事这方面工作，难点比较容易克服，而语音处理板的开发、声图转换技术的探索、整个系统的配合却是我们真正的难点。我们认为有如下几个关键技术要解决：

#### (1) 操作系统中设备驱动程序的剖析和修改

KX-600 设备驱动程序是由汇编语言写成的，其任务是把图形软件包送来的图形信息、文字信息和视频信息送到 KX-600 图形工作站。设备驱动程序根据给定的作图信息（功能、参数等）调用各个子模块把数据送到各个通道端口，显示处理机再把端口取到的数据送往显示屏显示。

TI-SPEECH 的驱动程序也是一个汇编语言接口软件，通过使用 DOS 的标准 I/O 调用访问语音硬件设备。

由于这两个驱动程序使用的端口地址发生了冲突，致使数据传输过程中发生错误。为此，需要分析驱动程序、修改驱动程序中有关的状态控制位，保证数据不会误传。

#### (2) MGP 图形软件包的修改

MGP 图形软件包适用于一般键盘作图，但对语音作图有时却无能为力。为此需要修改原有的图形软件包，建立语音参数设置模块，以实现语音识别、语音合成的初始化和装载语音算法程序。同时要建立语音识别、合成和透明键盘支持模块，以实现启动语音识别、语音合成功能，启动透明键盘支持软件，进行词表转换等功能。

#### (3) 透明键盘支持软件的研制

透明键盘支持软件是一个用汇编语言写成的系统程序，它把用语音输入的词组转换为对应的键盘内码，送到键盘缓冲区，从而实现了语音输入与键盘完全兼容。对用户来说，这一转换过程是透明的，用户可以同时进行语音和键盘输入；而不必考虑它们的内部过程。为此，在进行语音词表训练时，不仅要建立词组参考模板，还要建立每个词组对应的键盘内码。透明键盘支持软件在内存中为每个词表建立一个键盘码——词组转换对照表。

#### (4) 声图转换技术

声图转换是指把语音输入的词组转换成图形软件的调用格式。其最简单的方法是在计算机内存中建立一个声图对照表，但若增加一个词组，就需修改一次表的内容，很不方便。较好的方法是每个词组编写一个处理子程序。这种办法的好处是可以把标准图形软件包原封不动地移植过来，只要增加所有词组的处理程序即可完成声图转换，但是仍存在每增加一个词组就要为它编制一个处理程序的问题。

最好的办法是利用操作系统的键盘管理模块，把每个词组对应为一个键盘字符（或多个），每识别一个词组就把对应的键盘字符码送入 DOS 的键盘缓冲区。这样用户要增加一个词组，只要给出对应的键盘字符即可。

## 4 发展与展望

目前,计算机声图系统还是一个鲜为人们所涉猎的领域。一方面是因为语音识别还处于词组识别阶段,尽管国内外语音处理方面的专家都在自然语言理解方面做了大量的研究工作,但真正付诸实用的也只有少数几个专家系统。另一方面,由于图形系统的智能化主要是要有智能的高级语言支持,因此,人们的注意力集中到研制人工智能高级语言上。

然后,计算机语音输入毕竟是一种更为方便、自然和简单的输入手段。对用户来说,用语音输入比用任何一种其它输入设备都更为直观。因此,随着计算机图形技术和语音信号处理技术的发展,计算机声图系统将受到人们的关注。我们希望不久的将来能够出现具有分析推理能力的用自然语音输入的计算机图形系统。一种智能计算机图形系统模型如图 5 所示。

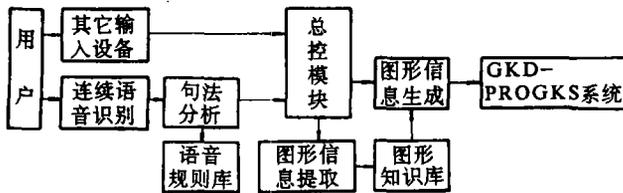


图 5 智能计算机图形系统模型

图中 GKD-PROGKS 图形系统是我校研制的、将 GKD-PROLOG 解释系统和国际图形标准软件 PLOT-10-GKS 结合起来的系统。由于采用了具有知识表达和知识获取能力的 PROLOG 语音,增强了命令解释、翻译等功能,因而使得整个系统具有简洁、方便和接近自然的特点。如果在此基础上,再扩大一些推理、演译模块,扩充自然语言处理和语音输入,就可产生如图 5 所示的智能化的计算机图形系统。

我们相信,经过科技人员不懈的努力,用人类自身的语言,同计算机对话,用语音命令来控制机器、控制作图的时代,不久就会到来。

### 参 考 文 献

- 1 空军三所,国防科技大学. MGP 军用图形包使用说明书. 1986
- 2 Flanagan. J. L Computer that talk and listen; Man-Machine Communication by voice. Proc IEEE 1976, 64 (1): 405~415
- 3 黄学东. 一个计算机汉字语音输入系统. 计算机学报, 1987, (6)
- 4 叶培建. 语言信号处理综述. 全国智能机器人, 语言识别, 计算机视觉会议论文集, 1987
- 5 俞铁诚. 通用实时语音识别系统. 物理学报, 1978, 27 (5)
- 6 牟迎春. GKD-PROGKS 图形系统的设计与实现. 国防科技大学硕士研究生论文, 1989

## A Computer Speech-graphics System Based on Speech Recognition

Xue Saoyu Liu Yucan Zhao Jinlong  
(Department of Electronic Technology)

(下转第 79 页)

### Abstract

This paper studies several problems for the cone ratio data envelopment analysis model; the existence of DEA effective decision-making units, the relations among several DEA efficiencies, the influence of input (output) cone and decision cone to DEA efficiency, and the properties of possible product set. The conclusions concerned have been obtained.

**Key words** decision making, model, data envelopment analysis, cone ratio, DEA efficiency

---

(上接 73 页)

### Abstract

A model of computer speech-graphics system is proposed in this paper. The paper describes its principle and its design method, and a simple computer speech-graphics system based on this method is developed. It consists of a TI-SPEECH speech processing system, a KX-600 graphics display system and an IBM-PC/AT computer. The graphics software package is MGP 2D military graphics package which is implemented in KX-600 graphics display system. It possesses the Transparent keyboard ability to support speech input. Finally, the paper proposes an intelligent computer graphics display model which consists of speech input, and GKD-PROGKS graphics system. It provides some valuable recommendations for the development of intelligent computer graphics display system.

**Key words** speech signal processing, display, microcomputer, transparent keyboard, speech-graphics system