

# 利用微机产生彩色语图的方法

孙金城 吕士楠

(中国科学院声学研究所)

1988年1月12日收到

本文在参考现有国内外计算机作语图的方法的基础上,提出用微机产生语图的方法.在不增加任何设备的条件下,利用软件作出彩色三维语图(没有彩色显示器时可作出黑白语图),质量可以和语图仪作出的语图相比拟。

## 一、引言

语图仪问世已有四十年的历史,至今仍是国内外语音实验室的重要研究工具.语图把语音这一复杂的时间过程以三维图谱的形式表示出来.即横轴表示时间,纵轴表示频率,用时间-频率平面上的灰度表示强度.语图使语言流中的各音素,包括元音,辅音及它们之间的联接形式表示得十分清楚.因而训练有素的专家可以从语图中辨认出说的是什么话.通过对语图的研究,解决了许多语音学中的基本问题.语图也用于许多实际工作领域,如通讯、语言训练、喉病诊断、发音人辨认、动物行为分析……等等.它的缺点是语图仪和专用语图纸价格昂贵、做图时间较长(每张图2.1s信号,需时3—5min),且不连续.尤其是语图难于给出谱的具体量值。

计算机和数字信号处理技术的发展使用计算机作语图成为可能.利用快速傅氏变换或数字滤波技术就可以得到如同语图仪三维分析的结果.问题在于表示方法.一个数字谱矩阵给出不同时间和不同频率的功率谱的准确值,用这些数据可以绘成三维的透视图.但这样给出的信息太多,不用说几千个谱值,就是三维透视图也已使人眼花缭乱,把握不住主要信息.国内已有一些用计算机作语图的研究结果发表<sup>[1-3]</sup>.国外的语音实验室大多也配有计算机

作语图的软件,但作图方法和国内文章介绍的方法不同.一般是在单色显示器上用亮点密度来表示谱值大小.具体地说,每一个时间和频率样点的谱值和屏幕上一个与表示字符相似的点阵相对应.谱值高,点阵中的亮点多,反之亦然.亮点在点阵中的分布可以采取各种形式,以达到理想的效果.西德科隆语音实验室采用随机分布形式,而日本名古屋大学板仓实验室则采用中心密集式.我们综合各种语图表示形式的优点,提出改进的中心密集式点阵的显示方法.主要改进有两点:1.在时间维和频率维上让两点阵之间都有重叠,以达到平滑的效果.2.对亮点进行彩色调制.下面具体描述显示方法和软件设计。

## 二、窄带语图显示

我们选用目前国内最流行的IBM-PC型微机为显示语图的工具.基本配置:IBM-PC/XT(或兼容机)、彩色显示器、彩色打印机、A/D(12bit)采样板.彩色显示器和图形板的分辨率越高则可获得更好的图形质量,但价格要贵.文中彩色语图是IBM-PC机为高分辨率(640×400)时所做。

显示屏幕用软件控制时只有两个亮度级可以利用,用单一光点表示不了谱值的大范围变化,而亮点密度可以在较大范围内变化.语言的动态范围一般在60dB以下,若窄带语图采用

× 5 点阵来表示谱级, 可以表示 25 种谱级。以每级 2dB 计算, 大致可以满足语言谱的要求, 精度已超过语图仪给出的灰度级的精度。下面以 5 × 5 点阵为例说明中心密集式亮点密度表示谱级的方法。表 1 中心点的显示阈值最低 (30dB)。

表 1 显示窄带语图用的 5 × 5 点阵中各点的显示阈值(dB)

| y | x  |    |    |    |    |
|---|----|----|----|----|----|
|   | 1  | 2  | 3  | 4  | 5  |
| 1 | 74 | 68 | 54 | 70 | 72 |
| 2 | 60 | 42 | 38 | 40 | 58 |
| 3 | 50 | 34 | 30 | 32 | 48 |
| 4 | 62 | 44 | 36 | 46 | 56 |
| 5 | 76 | 66 | 52 | 64 | 78 |

假设 30dB 以下无显示, 达到或超过 30dB 则(3, 3)点亮,(4,3)接近中心,显示阈值 32dB。接着是(2, 3)、(3, 4)、(3, 2)……离中心越来越远,每递进一点,顺序使阈值提高 2dB。达到或

超过 78dB 时第(5, 5)点才亮。谱值越高,亮点越多。这是板仓语音实验室使用的方法。我们在此基础上又规定谱值每隔 10dB 左右改变一种颜色,红色谱值最高,黄色次之,绿色再次之。并且使点阵在时间维和频率维每次只向前推进两个点,即有三个点是重叠的,以达到时间维和频率维平滑的目的,使窄带语图呈现出清晰的谐波结构。

图 1 是用上述方法作出的“他去无锡”四个音节的语图。原始数据采样率为 8 kHz, 帧长 128 点,即 16ms; 用 0 补足到 256 点,加海明窗后进行快速傅氏变换;等效带宽约 60Hz。从图 1 中可已看到“他”是阴平声,谐波曲线构成一组水平线;“去”是去声,属降调,谐波曲线向下倾斜;“无”是阳平,曲线向上倾斜。由于“无(wu)”由一个半元音开始,使“无”和“去”的联结非常紧密,声调变化也是连续的。因此谐波曲线呈现为一马鞍形。结果与用语图仪作出的 45Hz 带宽的语图非常接近。

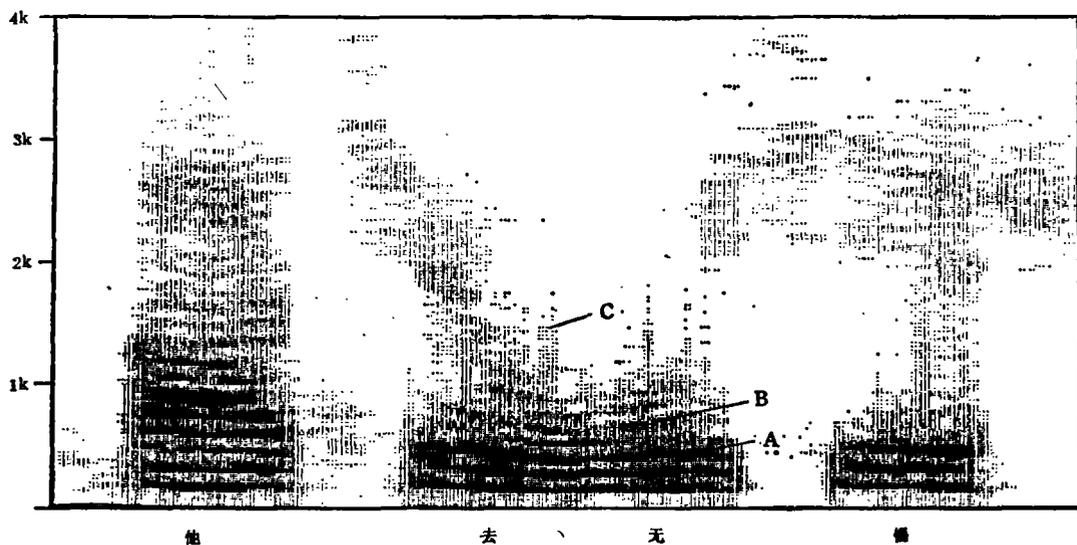


图 1 IBM-PC 作窄带语图

注: 图 1 原是彩色图,因故改为黑白图,图中 A 类是红色, B 类是黄色, C 类是绿色。

### 三、宽带语图显示

和窄带语图不同, 宽带语图主要用于显示

语音的共振峰结构。调整 FFT 实际占有数据部分的长度, 可以灵活地改变分析的等效带宽。在作宽带语图时, 由于实际帧长缩短, 使语图的时间轴展宽, 因而要对点阵作调整, 并使其在时

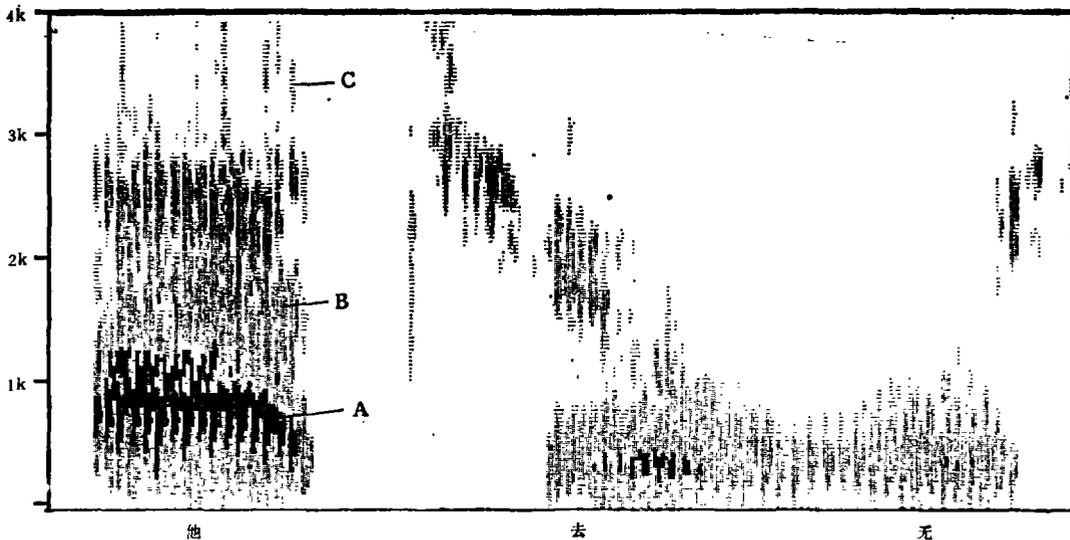


图2 IBM-PC 作宽带语图

注：图2原是彩色图，因故改为黑白图，图中A类是红色，B类是黄色，C类是绿色。

间轴上点阵不相重合，以提高时间维的分辨力，显示出声门脉冲的间隔。

计算机显示“他去无锡市”中的“他去无”三个音节的宽带语图见图2。它的帧长为32点（补零扩展成256点后作FFT），每帧推移量为16点，信号的实际帧长为4ms，等效分析带宽约250Hz。从图2中可以看到在截止频率4000Hz以下一般有三到四个共振峰。如“他”的第一、二、三共振峰频率分别为1000Hz、1100Hz、2500Hz左右。由于协同发音影响，第一、三峰向下弯，第二峰向上弯，与300Hz带宽的语图仪所作的语图结果完全一致。

#### 四、软件设计

快速傅氏变换是用北方交大提供的FFT软件进行的。语音信号经低通滤波和采样后分段递推进行傅氏变换，取模、取对数后以谱矩阵形式存入磁盘，显示图形时再从磁盘读出数据。作图程序用BASIC语言编写，为加快运行速度，要用87BASCOM编译器编译BASIC源程序。完成一幅占满屏幕的语图约120s。

#### 五、结束语

微计算机作彩色语图的特点：1. 色彩使三维语图的表示更直观明了，便与把握住语音中最主要的内容。最多可以用16种颜色显示16个强度等级的谱值，但颜色过多反而使人眼花缭乱，抓不住语音中的主要信息。如遇特殊需要，稍加几条语句便能获得多颜色多层次的强度维显示。2. 动态范围易于调整，只要改变语音频谱的阈值，即可获得不同动态范围的语图。3. 频率范围视采样频率而定，可有多种上限频率。4. 等效滤波器带宽视FFT样点而定，易于调整 and 实现。5. 可以人机对话方式编辑屏幕图形，如可利用色彩醒目的语音基频、谐波或共振峰的条纹，方便地在显示屏幕上搬挪，拼接已被打乱的语音信号，恢复原始语音的语图。6. 解决了语图仪做定量分析的困难。7. 图形连续，操作简单，价格便宜，便于推广。

本方法尚有不足之处：产生的语图还没有语图仪所作语图那么精细，尚需进一步改进。

## 参 考 文 献

- [1] 滕文善,应用声学,3-3(1984),24.  
[2] 包紫薇、华一满,应用声学,4-1(1985),12.

- [3] Zhang Libe & Zheng Yi, The Realtime Pseudocolor Enchoding of Sound Spectrograms, Advance in Signal Processing, The Proc. of 1986 IEEE-Academia Sinica Workshorp on ASSP. p 253—256.

# 杭州电声厂混响室的声学设计和鉴测

项端祈 赵一兴 王峰 陈金京 葛砚刚

(北京市建筑设计院研究所)

1987年10月26日收到

使混响室获得扩散声场是声学设计中最主要的问题。在我国已建的混响室建筑中,虽然都做了各种为实现声扩散的努力,但能获得预期效果的不多,从而不得不追加扩散结构。电声厂混响室用不规则室形和球切面扩散结构相结合的方法,获得了良好声扩散效果。本文将着重介绍有关声扩散的设计,同时,对该混响室的声学设计和鉴测也作概要的叙述。

## 一、概 述

杭州电声厂混响室是该厂研究所声学实验楼(它还有消声室、测听室和其它声学实验的测试中心楼)的一个组成部分。混响室的功能很多:可测定材料的吸声系数,扬声器的效率,声源的声功率和频谱,传声器的互易校准,空气中的声吸收;对灵敏机件作耐噪声试验(噪声疲劳)和产生人工混响等。电声厂混响室主要作前三项试验用。

混响室平面为接近矩形的不规则形,有效容积 $268.6\text{m}^3$ ,总表面积为 $274.8\text{m}^2$ 。混响室的平、剖面见图1所示。

混响室的声学设计指标经商定如下:

**混响时间:**原则上应尽可能加长室内混响时间,但最低限度应满足如下要求:

$$125\text{Hz} \leq 15\text{s}; 500\text{Hz} \leq 10\text{s};$$

高频混响应达到高限  $T_g \cong \frac{2s}{f}$  ( $f$ 为千赫)的一半以上,即 $2000\text{Hz} \leq 6.0\text{s}$ 。<sup>[1]</sup>

**声场不均匀度** 在测试区内,各点最大与最小声压级的差值  $\Delta P = P_{\max} - P_{\min}$  在低频应

小于6dB,中高频应低于3dB,即: $125\text{Hz} \leq \pm 3.0\text{dB}; 500 \sim 2000\text{Hz} \leq \pm 1.5\text{dB}$ 。

**声扩散** 无定量指标,要求扩散值 $d$ (传声器水平面的方向特性)尽可能大,同时要求在不追加扩散结构的条件下能获得良好的声扩散。

**室内允许噪声级** 要求在任何时刻,在周围噪声和振动的同时作用下,室内噪声级 $\leq 30\text{dBA}$ 。

为了达到上述各项声学指标,在调查研究的基础上进行设计,并注意到施工各阶段可能出现的问题,从而使建成后的混响室,通过声学鉴测,达到了预期的效果。有的声学指标比预计的好。

## 二、混响室的声学设计

### 1. 混响室容积的确定

为使混响室获得扩散声场,就须在测试频带(100Hz~4000Hz)宽度内激发室内足够多的简正方式(主要在低频范围内)。在频率 $f$ 附近带宽 $\Delta f$ 内的简正振动方式的数目 $\Delta N$ 可由下式求得<sup>[2]</sup>: