

全息声谱显示

朱榕成 张礼和

(杭州大学电子工程系)

1990年11月15日收到

本文提出了一种全息声谱的编码和显示方法。全息声谱图是在二维屏幕上显示四维的声谱函数语图。它同时表征了语声的幅度谱和相位谱。在显示坐标上,除了时间轴和频率轴两维外,还分别用亮度来表示幅度谱,用色彩来表示相位谱。通过编码,我们成功地实现了这一显示系统。

一、引言

众所周知,傅立叶谱在包括信号处理的众多领域里有着广泛的用途;而对语声信号的频域分析和显示则是语声信号研究的一种重要技术手段。在传统的声谱的语图显示方式上,通常用平面本身的两维坐标来表示时间和频率,而用曲线的灰度等级来表示功率谱(幅度谱)值的大小^[1-7]。然而这种语图所阐述的信息量目前已显得贫乏。近年来的大量理论和实验研究表明,信号的相位谱同样包含着众多的信息;甚至在相当意义上说,信号(包括语声信号)的相位谱蕴含的信息量比信号的幅度谱所含的信息量还要大^[8-11]。这自然引起了人们对信号相位谱的兴趣以及对其进行更深入的研究的热情,也提出了对相位谱直观显示的要求。在诸如通信的其它许多领域,相位谱都起着重要的作用。我们提出的全息声谱的概念也就是源于这样的要求。

不同于幅度谱语图的是,全息声谱图能在一张平面语图上同时显示出信号的幅度谱和相位谱,由此大大提高了语图所表征的信号的信息量。但这也提出了一个相当的难题,它要求我们在一张二维的平面上来表示四维矢量。采用通常的方法显然是无法做到的,在研究了各种可行的表征方式之后^[12],我们选择了用亮度来表示信号的幅度谱值,用色彩来表示信号的

相位谱值。而其它两维仍然是时间维和频率维。因此,这样的全息声谱图上就可以通过亮度的强弱来表示出信号能量的大小,通过颜色的变化来看出相位的变化。我们通过一台微机和一套图象显示系统,成功地实现了这一目的^[13]。其结果表明,全息声谱图具有良好的可读性,很高的准确度以及相当的灵活性。它可以广泛地应用于语声分析,语音识别,临床喉音诊断,布料配色,公安侦察等众多的领域。

二、离散傅立叶谱的生成

首先,我们假设有一连续的模拟信号 $s(t)$, 经 A/D 取样后,成为一离散信号 $S(k)$, $k = 0, 1, 2, \dots$ 。

考虑一段长度为 N 的时域信号 $S(k)$, 其第 l 帧离散傅立叶变换 (DFT) 为:

$$F_l(n) = \sum_{k=0}^{N-1} s_l(k) w(k) W_N^{nk} \\ n = 0, 1, \dots, N-1 \\ l = 0, 1, \dots, M-1 \quad (1)$$

其中 $W_N = \exp(-j2\pi/N)$; $w(\cdot)$ 为某个窗函数,可按需要加以选择; N 为帧数。为简便起见,以下把 $F_l(n)$ 简记为 $F(n)$, $S_l(n)$ 简记为 $S(n)$ 。这样, $F(n)$ 就可表示为

$$F(n) = F_r(n) + jF_i(n) = |F(n)|e^{j\varphi(n)} \quad (2)$$

其中 $|F(n)|$ 为幅度值, $\varphi(n)$ 为相位谱; $F_r(n)$, $F_i(n)$ 分别为谱的实部和虚部。

对全息声谱编码时,幅度值需进行归一化:

$$|F(n)| = |F(n)|/F_m \quad (3)$$

其中 F_m 是所有的幅值中的最大值,亦即

$$F_m = \max\{|F_i(n)|\} \quad (4)$$

相位取主值为

$$\begin{aligned} \varphi(n) &= \arg[F(n)] \\ 0 &\leq \varphi(n) < 2\pi \end{aligned}$$

相位也需要进行规正,以便能显示标准化的相位,便于显示时观察。文献[13]对此有较详细的论述。

三、全息声谱的显示

全息声谱显示的一个关键问题是,怎样在现有的实验室条件下,在二维显示屏上来表示四维的全息声谱矢量。经过综合分析和研究^[12-13],我们最终选定了用亮度和色彩作为除了时间和频率二维外的附加二维的坐标分量。这样,一个全息声谱的矢量,包含着时间、频率、幅度谱和相位谱四个变量,是四维函数。它可以在二维的彩色显示屏幕上表示出来。实验证明这是可行的。

假设全息声谱为 $\vec{F}(n)$, 将其分解为三基色 $\vec{R}(n)$, $\vec{G}(n)$, $\vec{B}(n)$ 分别代表红、绿、蓝三基色分量,并简记为 $\vec{R}$, $\vec{G}$, $\vec{B}$ 。则

$$\begin{aligned} \vec{F}(n) &= \vec{R} + \vec{G} + \vec{B} \\ &= R\vec{R}_0 + G\vec{G}_0 + B\vec{B}_0 \end{aligned} \quad (5)$$

其中 $\vec{R}_0$, $\vec{G}_0$, $\vec{B}_0$ 分别是 $\vec{R}$, $\vec{G}$, $\vec{B}$ 的单位矢量。

下面,我们要找出满足 $\vec{F}(n)$, 或

$$F(n) = |F(n)|e^{j\varphi(n)}$$

的条件下的 R, G, B 三分量之值。现用 $b(n)$ 表示亮度值。显然,基于等亮度的要求,存在以下关系式:

$$R + G + B = f(|F(n)|) = b(n) \quad (6)$$

而 R, G, B 三分量具体的选值,则依赖于编码系统的建立。

一般地,可以取 $f(|F(n)|) = |F(n)|$, 但这常常需要根据编码系统的实际情况而加以选

应用声学

择。

对于色彩 $C(n)$ 而言, $C(n)$ 是 R, G, B 的函数;即

$$C(n) = C(R, G, B)$$

而 R, G, B 又分别是 $|F(n)|, \varphi(n)$ 的函数,即

$$\begin{cases} R(n) = R(|F(n)|, \varphi(n)) \\ G(n) = G(|F(n)|, \varphi(n)) \\ B(n) = B(|F(n)|, \varphi(n)) \end{cases} \quad (7)$$

因此, $C(n)$ 是 $|F(n)|, \varphi(n)$ 的函数,即

$$C(n) = C(|F(n)|, \varphi(n))$$

我们的目的是希望 $C(n)$ 仅是 $\varphi(n)$ 的函数,而不依赖于 $|F(n)|$ 而变,即

$$C(n) = C(\varphi(n))$$

由于在式(6)中,有 $R, G, B = g_i(|F(n)|)$ 的满足 [$g_i(n)$ 是某个函数, i 代表 R, G, B], 故式(7)可演变为

$$\begin{cases} R(n) = R(\varphi(n)) \\ G(n) = G(\varphi(n)) \\ B(n) = B(\varphi(n)) \end{cases} \quad (8)$$

同时,使式(8)满足相位谱周期性的要求,以避免编码时色彩的非连续性,这样有下式

$$\begin{cases} R(n) = R(\varphi(n) + 2\pi K) \\ G(n) = G(\varphi(n) + 2\pi K) \\ B(n) = B(\varphi(n) + 2\pi K) \end{cases} \quad K \in J \quad (9)$$

J 为整数集合。根据式(6)和(8)建立起的系统,就实现了声信号的全息频谱显示。这种编码的结果是,使亮度表示幅度谱,使色彩表示相位谱。

四、全息声谱的编码

本节讨论对幅度谱和相位谱同时进行编码的全息声谱的实现方法。复傅立叶谱可写成

$$F(n) = F_r(n) + jF_i(n)$$

或

$$F(n) = |F(n)|e^{j\varphi(n)}$$

为了表达方便,我们把上式写成二维矢量的形式:

$$\vec{F}(n) = \vec{F}_r(n) + j\vec{F}_i(n) \quad (10)$$

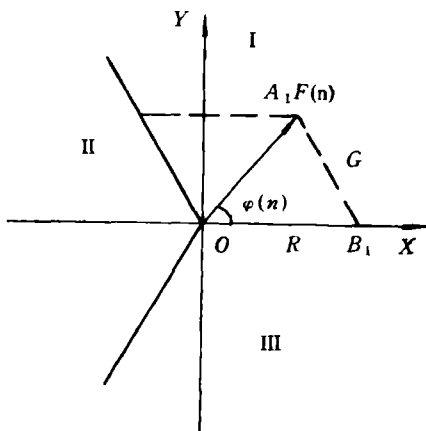


图1 编码平面

$$\text{或} \quad \vec{F(n)} = F_r(n)\vec{x}_0 + F_i(n)\vec{y}_0 \quad (11)$$

其中 $\vec{x}_0, \vec{y}_0$ 分别是 $\vec{X}, \vec{Y}$ 的单位矢量。

把 $X-Y$ 平面均分为三个独立的部分: I 区, II 区, III 区;在平面上的每个矢量都可分解为色彩矢量 $\vec{R}, \vec{G}$, 和 $\vec{B}$ (如图 1 所示)。 $\vec{R}_0, \vec{G}_0, \vec{B}_0$ 分别表示 $\vec{R}, \vec{G}, \vec{B}$ 的单位矢量。因此,当

$$0 \leq \varphi(n) < \frac{2}{3}\pi$$

时, $\vec{F(n)}$ 在 I 区;当

$$\frac{2}{3}\pi \leq \varphi(n) < \frac{4}{3}\pi$$

时, $\vec{F(n)}$ 在 II 区;当

$$\frac{4}{3}\pi \leq \varphi(n) < 2\pi$$

时, $\vec{F(n)}$ 在 III 区。

显然,平面上的色彩分布是连续的,由于某种原因(如计算误差)而引起的相位的边缘(在 0 和 2π 处)色彩“跳动”影响就可消去。

首先,暂不考虑亮度,在色彩分量上对频谱进行编码。在 I 区,根据正弦定律,在 $\triangle OA_1B_1$ 中, R, G 和 $|F(n)|$ 满足下列关系:

$$\frac{R}{\sin\left[\varphi_1(n) + \frac{\pi}{3}\right]} = \frac{G}{\sin\varphi_1(n)}$$

$$= \frac{|F(n)|}{\sin\frac{\pi}{3}} \quad (12)$$

式中 $\varphi_1(n) = \varphi(n)$ 。使

$$|\widetilde{F(n)}| = |F(n)| / \sin\frac{\pi}{3} \quad (13)$$

$$\text{则有} \quad R = |\widetilde{F(n)}| \sin\left[\varphi_1(n) + \frac{\pi}{3}\right]$$

$$G = |\widetilde{F(n)}| \sin\varphi_1(n) \quad (14)$$

在色彩编码完成后,现在考虑亮度。为了满足等亮度条件式(6),我们设

$$b(n) = 2|F(n)| = \sqrt{3}|\widetilde{F(n)}|$$

$$\text{则} \quad B = \sqrt{3}|\widetilde{F(n)}| - (R + G)$$

这样,在 I 区的三个分量为

$$\begin{cases} R = |\widetilde{F(n)}| \sin\left[\varphi_1(n) + \frac{\pi}{3}\right] \\ G = |\widetilde{F(n)}| \sin\varphi_1(n) \\ B = \sqrt{3}|\widetilde{F(n)}| - (R + G) \end{cases} \quad (15)$$

其中 $\varphi_1(n) = \varphi(n)$ 。

同样地,在 II 区,可得结果

$$\begin{cases} G = |\widetilde{F(n)}| \sin\left[\varphi_{II}(n) + \frac{\pi}{3}\right] \\ B = |\widetilde{F(n)}| \sin\varphi_{II}(n) \\ R = \sqrt{3}|\widetilde{F(n)}| - (G + B) \end{cases} \quad (16)$$

$$\text{其中} \quad \varphi_{II}(n) = \varphi(n) - \frac{2\pi}{3}.$$

在 III 区,有:

$$\begin{cases} B = |\widetilde{F(n)}| \sin\left[\varphi_{III}(n) + \frac{\pi}{3}\right] \\ R = |\widetilde{F(n)}| \sin\varphi_{III}(n) \\ G = \sqrt{3}|\widetilde{F(n)}| - (B + R) \end{cases} \quad (17)$$

$$\text{式中} \quad \varphi_{III}(n) = \varphi(n) - \frac{4\pi}{3}.$$

通过上述过程,就完成了全息声谱在二维平面上的编码。特别地,当我们把亮度设为常数时,就可得到相位谱;而当我们把 R, G, B 等值归一化时,就得到了幅度谱。

五、全息声谱显示系统及实验结果

基于上述编码和显示原理,我们使用了一

台微机以及配套的图象显示系统成功地实现了全息声谱显示系统.本系统所产生的全息谱图,其最长时间区间为 10s, 频率范围为 80Hz—4.5KHz, 频率分辨率约为 35Hz. 语言信号经放大/衰减接口,输入动态范围可调.幅值量化精度为 12bit, 加权窗函数为矩形窗、海宁窗、汉明窗,升余弦窗等(可任选).数据所需存贮空间约为 150K 字节.

作为显示实例,图 2 给出了标准汉语中的四个声母(浊音) {a,i,u,o} 的全息声谱图.从图中可以明显地看出各浊音幅谱的强弱,也可以清晰地分辨出不同音素的相位谱值.同时,我们还可以容易地将音素从背景噪声中分辨出来.这种全方位的,直观的语图显示,为声学研究的深入开展提供了一种技术手段.

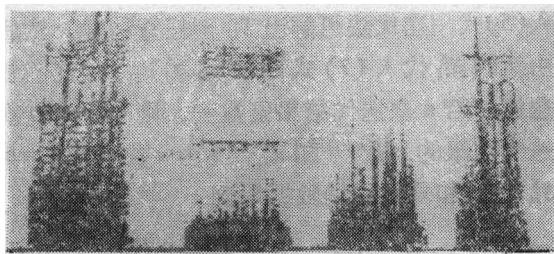


图 2 {a,i,u,o} 全息声谱图(色彩已黑白化)

全息声谱图不仅可以显示观察,同时也提供了定量的数值.在整套显示系统中,多功能和灵活的编程使得全息声谱显示的应用更趋方便、实用和广泛.

参 考 文 献

- [1] Koenig W. et al., *J. Acoust. Soc. Amer.*, **18-7** (1946), 19—49
- [2] Oppenheim A. V., *IEEE Spectrum*, **7-8** (1970), 57—62
- [3] Mermelstein, P. *IEEE Trans. Audio Electroacoust.*, **19** (1971), 44—47
- [4] Buchanan M. D. et al., *Electro-Optical Systems Design*, **12-3** (1980), 29—36
- [5] 包紫薇,华一满,应用声学, **1**(1985),12—16
- [6] 滕文善,声学学报, **1**(1986),56—60
- [7] Zhang Lihe and Zheng Yi, *Proc. WASSP Conf.*, 1986, 253—256
- [8] Oppenheim A. V. and Lim, J. S. *Proc. IEEE*, **69-5** (1981), 529—541
- [9] Hayes M. H. et al., *IEEE Trans. ASSP*, **28-12** (1980), 672—680
- [10] Hayes M. H., *IEEE Trans. ASSP*, **30-4** (1982), 140—154
- [11] Lane R. G. et al., *IEEE Trans. ASSP*, **35-4** (1987), 520—525
- [12] 朱榕成等,中国微机七届年会论文集,1990,106—110
- [13] Zhu Rongcheng et al., *Proc. 3rd Western/Pacific Regional Acoust. Conf.*, 1988, 191—194

PVDF 压电聚合物复常数的测量及其温度特性

朱 厚 卿 张 金 铎

(中国科学院声学所) (北京大学无线电系)

1990 年 8 月 6 日收到

本文报道了根据测得的 PVDF 压电薄膜横向长度伸缩振动的导纳圆图,使用 Smits^[1] 方法,精确测定了压电薄膜的复介电、压电和弹性常数及其温度特性.在 17℃—110℃ 的温度范围内,介电常数的实部为 10—13,其虚部与实部之比小于 5%;压电常数的实部随温度上升而增加,在 85℃ 附近有一极大值, $d_{31} = 27 \times 10^{-12} \text{C/N}$. 超过 100℃ 的 d_{31} 急剧下降.

一、引 言

PVDF 压电聚合物是一种发展较快的新型

应用声学

敏感和换能材料,其特点是有从甚低频到千兆赫的宽频率范围;有大的动态范围,从 10^{-5} 到 10^8N/m^2 ;有很低的声阻抗和很低的机械 Q 值;有高的压电效应和大的热释电效应.它正在并