

用白光光学处理技术实现 语音的频谱分析

杨震寰 林聪能

(美国宾夕法尼亚州州立大学电机系)

许克彬

(哈尔滨工业大学物理教研室)

提 要

本文提出了利用白光光学频谱分析系统获得语音信号彩色编码频谱的技术。借助于阴极射线示波器,首先将录制在磁带上的语音信号用密度调制的方法转变为照相底片的振幅透射率的变化,从而实现信号由时域到空域的转变。然后在白光光学频谱分析系统中,利用放置在傅里叶频谱面上的取样狭缝得到按频率假彩色编码的频谱图。文中用部分相干理论对白光光学频谱分析系统作了初步的讨论,并给出了若干中文语音信号的频率假彩色编码谱图。

一、引 言

语音信号的频谱分析作为一种基本的手段广泛地应用于语言学、语音学、语病学、语言心理学、语音的合成和识别中。大约40年前,贝尔电话实验室的W. Koenig等人^[1]首次用电子学方法实现了各种声音信号的频谱分析,至今利用电子学技术在语音合成和识别方面已取得很大进展。Cutrona等^[2]最早提出多通道光学频谱分析器的设想,其后Thomas^[3]建议了处理宽带信号的频谱分析系统。Alexoff利用相干光学系统得到了第一个带有音乐背景的语音谱图^[4]。

本文提供一个用白光光学频谱分析系统进行语音信号频谱分析的方法。

二、语音信号从时域到空域的转变

由于光学频谱分析是一种空间处理,而语音信号是时间信号,因此首先必须完成语音信号由时域到空域的转变,密度调制^[5]则是实现这一转变的最直接和常用的技术。

将预先录制在磁带上的语音信号通过录音机的电压输出馈送到阴极射线示波器(CRT)的Z轴,用以调制其光斑的强度和变化频率,并保持光斑强度的变化与馈送信号的幅度成正比(由于馈送电压的双极性,一个适宜选择的直流偏置是必要的)。调整示波器垂直方向的扫描频率使其远高于语音信号中的高频分量,垂向扫描的结果使示波器的屏上呈现一强

度均匀的垂直亮带,其强度由语音信号幅度线性调制,其变化频率则与语音信号同步。将此亮带信号用一水平扫描的照相装置记录在底片上,实验装置如图 1(a)所示。经过处理后的空间信号片的振幅透射率应与示波器上被调制的亮带的强度成线性关系。此类光栅状的空间信号片如图 1(b)所示,即可在光学频谱分析系统中进行处理。

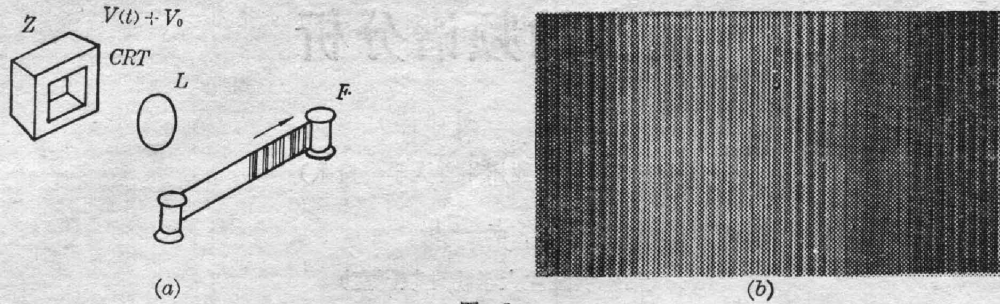


图 1

(a) 把时间信号转变为空间模式的装置; (b) 典型的空间信号片

Fig. 1

(a) Arrangement for converting a time signal into a spatial format;

(b) A typical spatial signal format

当扫描照相装置的水平扫描速度为 v 时,语音信号的时间频率分量 ν 与空间信号片的空间频率分量 u 之间的对应关系如下:

$$u = \nu / v. \quad (1)$$

在理想情况下,空间信号片的高频响应极限应由底片的分辨率决定,实际上它受制于示波器光斑的有限尺寸。典型的语音信号的带宽大约是 4 kHz,而且其谐波分量的振幅通常随频率的增加而减小。为了减少高频信息损失,应适当提高扫描照相装置的水平扫描速度。此外,当信号馈送到示波器 Z 轴之前,应先经过一放大倍率随信号频率增加而增大的“预加强电路”(pre-emphasis circuit)。

空间信号片的有效低频响应极限由后继的光学频谱分析系统决定,即主要由光源尺寸及输入平面的窗口尺寸决定。当光源很小时,如果输入平面上的窗口宽度为 W_F ,取有效的低频响应极限 $u_L \cong 2/W_F$,于是扫描照相装置的水平扫描速度应满足如下关系: $v \leq W_F \nu / 2$ 。因此,一个适宜的水平扫描速度的确定将是上述考虑的折衷产物。

三、白光光学频谱分析系统

在图 2 所示的白光光学频谱分析系统中,位于输入平面 P_1 上的空间信号片的线条平行于 x 轴,光栅 G 的刻线平行于 y 轴,从而在傅里叶频谱面 P_2 上得到输入空间信号片的各级辐射状彩虹谱图。放置在傅里叶频谱面 P_2 上具有取样狭缝的模板,只截取某一级彩虹谱图,使不同频率的谱显示不同颜色,如图 3 所示。(为了使取样狭缝保持垂直位置,光栅及输入空间信号片成 θ 角的倾斜,见图 2。)透镜 L_3 使取样后的频谱成像于输出平面 P_3 上,放置于平面 P_3 上的记录底片与输入空间信号片同步地连续运动,因此记录下全部信号频谱。

上述白光光学频谱分析系统是部分相干的。假定色散作用与傅里叶变换作用相互独立,为简单起见,设光源为方形均匀源,边长为 b 。考虑一级衍射,等效于光栅把光源沿 x 方

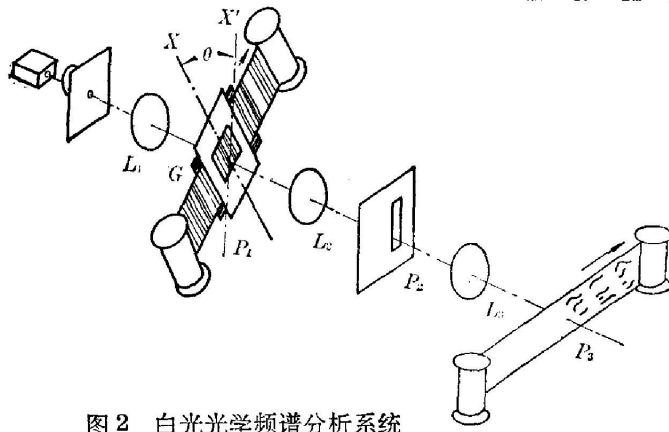


图2 白光光学频谱分析系统

Fig. 2 White-light optical spectrum analyzer

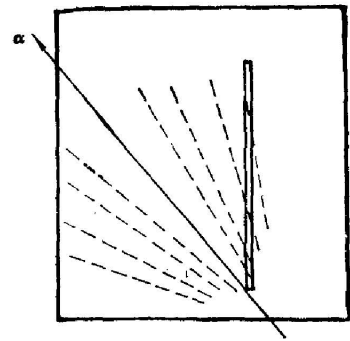


图3 傅里叶彩虹频谱的取样

Fig. 3 Sampling Fourier rainbow color spectrum

向色散成一条宽度为 b 的彩虹带, 其强度分布可近似表示为

$$\gamma(x_0, y_0; \lambda) = AS(\lambda) \text{rect}[(x_0 + u_0 \lambda f_1)/b] \text{rect}(y_0/b), \quad (2)$$

式中 A 为一常数, $S(\lambda)$ 为照明光源的光谱分布, u_0 为光栅频率, f_1 为准直透镜 L_1 的焦距, (x_0, y_0) 为光源平面坐标。

上述照明条件在输入平面 P_1 上引起的互强度函数为(见图 4)。

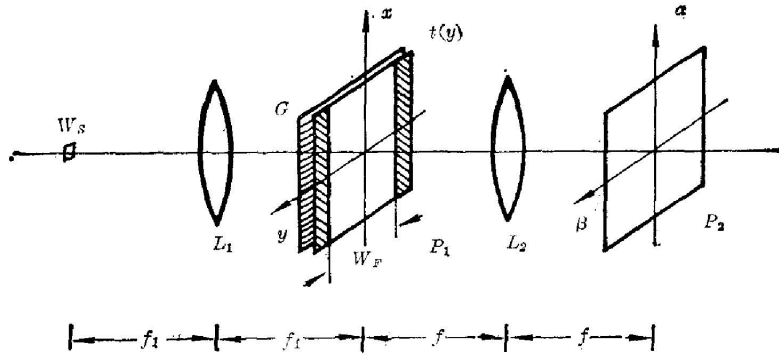


图4 实验装置示意图

Fig. 4 Schematic diagram of the experimental arrangement

$$I(x_1, y_1; x_2, y_2; \lambda) = \iint_{-\infty}^{\infty} \gamma(x_0, y_0; \lambda) \cdot \exp\{(-i2\pi/\lambda f_1)[x_0(x_1 - x_2) + y_0(y_1 - y_2)]\} dx_0 dy_0 \quad (3)$$

或

$$I(x_1, y_1; x_2, y_2; \lambda) = AS(\lambda) b^2 \text{sinc}[b(x_1 - x_2)/\lambda f_1] \cdot \exp[i2\pi u_0(x_1 - x_2)] \cdot \text{sinc}[b(y_1 - y_2)/\lambda f_1]. \quad (4)$$

如果空间信号片的复数振幅透射率为 $t(y)$, 输入平面上的窗口函数为 $\text{rect}(y/W_F)$, 则在空间信号片后面的互强度函数为

$$I(x_1, y_1; x_2, y_2; \lambda) = I(x_1, y_1; x_2, y_2; \lambda) t(y_1) t^*(y_2) \text{rect}(y_1/W_F) \text{rect}(y_2/W_F). \quad (5)$$

由互强度函数的传播定理, 在傅里叶频谱面 P_2 上

$$I(p_1, q_1; p_2, q_2; \lambda) = (c/\lambda f) \iiint_{-\infty}^{\infty} I(x_1, y_1; x_2, y_2; \lambda)$$

$$\cdot \exp[-i(x_1 p_1 + y_1 q_1 - x_2 p_2 - y_2 q_2)] dx_1 dy_1 dx_2 dy_2, \quad (6)$$

式中 $p = 2\pi\alpha/\lambda f$, $q = 2\pi\beta/\lambda f$ 。

将(4)、(5)式代入(6), 在 $p_1 = p_2 = p$, $q_1 = q_2 = q$ 处,

$$I(p, q; \lambda) = K S(\lambda) \operatorname{rect}\left[\lambda f_1 \left(\frac{p}{2\pi} - u_0\right) / b\right] \iint_{-\infty}^{\infty} \operatorname{sinc}[b(y_1 - y_2)/\lambda f_1] \\ \cdot t(y_1) t^*(y_2) \operatorname{rect}(y_1/W_F) \operatorname{rect}(y_2/W_F) \exp[-i(y_1 - y_2)q] dy_1 dy_2, \quad (7)$$

式中 K 为一常数。

为简便计, 取空间信号片的某一空间频率分量 u_1 为例, 令 $t(y_1) = 1 + m \cos 2\pi u_1 y_1$, $t^*(y_2) = 1 + m \cos 2\pi u_1 y_2$ ($m < 1$), 且只考虑一级衍射,

$$t(y_1) t^*(y_2) = (m^2/4) \exp[i 2\pi u_1 (y_1 - y_2)]. \quad (8)$$

将(8)式代入(6)式, 得

$$I(p, q; \lambda) = K' S(\lambda) \operatorname{rect}\left[\lambda f_1 \left(\frac{p}{2\pi} - u_0\right) / b\right] I(q; \lambda), \\ I(q; \lambda) = \iint_{-\infty}^{\infty} \operatorname{sinc}[b(y_1 - y_2)/\lambda f_1] \exp[i 2\pi u_1 (y_1 - y_2)] \operatorname{rect}(y_1/W_F) \\ \cdot \operatorname{rect}(y_2/W_F) \exp[-i(y_1 - y_2)q] dy_1 dy_2. \quad (9)$$

考虑一完全相干的极限情况, 此时 $b \rightarrow 0$,

$$\lim [b(y_1 - y_2)/\lambda f_1] = 1, I(q; \lambda) = W_F^2 \operatorname{sinc}^2 [W_F(q - 2\pi u_1)/2\pi].$$

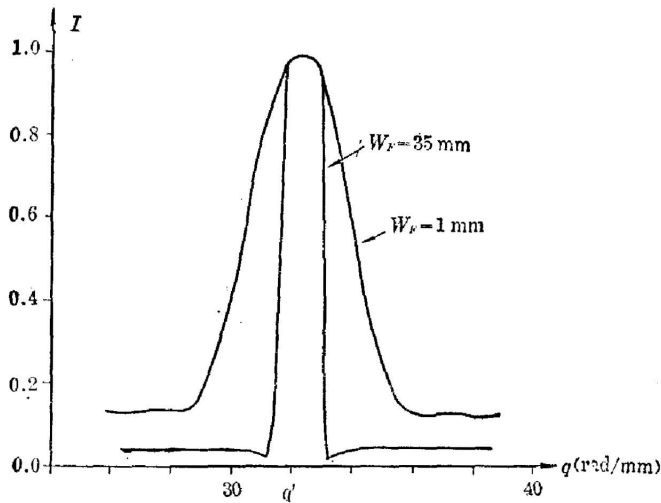


图 5 对两不同尺寸的输入窗口, 谱线的强度分布

Fig. 5 Intensity distribution of the spectral lines for two different values of the input window sizes

取 $\lambda = 0.55 \mu\text{m}$, $b = 90 \mu\text{m}$, $v = 195 \text{ mm/s}$, $u_1 \approx 5.13 \text{ cycle/mm}$ ($\nu_1 = 1000 \text{ Hz}$), $f_1 = 762 \text{ mm}$, $f = 360 \text{ mm}$, 可得归一化的函数 $I(q)$, 如图 5 所示, 曲线的中心频率即所选取的空间信号片的空间频率分量 $q' = 2\pi u_1$, 所对应的空间坐标为 $\beta' = \lambda f q' / 2\pi = \lambda f u_1$ 。由式 $I(p, q; \lambda)$ 可知在 p (或 α) 方向上, $p' = 2\pi u_0$, $\alpha' = \lambda f_2 u_0$ 。可见在频谱面的空间坐标系中, 谱线的中心位置 (α' , β') 沿 α , β 方向均正比于波长而变化。对空间信号片的空间频率分量 u_1 , 在白光照明下, 其频谱是斜率为 $f_1 u_0 /$

$f u_1$ 的彩虹带。基于彩虹带的斜率反比于信号片的空间频率分量, 考虑空间信号片的所有空间频率分量, 其频谱将是辐射状的彩虹谱图。

谱图的频率分辨率由光源的大小及输入窗口尺寸决定, 对于这两种独立增宽因素, 谱线的宽度 (以空间频率表示) 可写成 $\Delta q = 2\pi \sqrt{(b/\lambda f_1)^2 + (1/W_F)^2}$, 由于在空间频率平

面上以取样狭缝截取频谱,谱图表现的频率分辨率将与取样狭缝的宽度和取向密切相关。对较小的光源尺寸,改变输入平面上窗口 W_F 的大小,可以得到宽带或窄带谱图。

四、实验结果

在语音信号的时空转换装置中,我们选取 Technical Pan Film 2415 作为记录底片,由于它有较高的感光灵敏度和分辨本领,而且对比度可以在处理过程中调整。在白光光学频谱分析系统中,用 75 W 的 ORIEL 氙灯照明直径 $90\ \mu\text{m}$ 的针孔作为白光光源, $80\ \text{line/mm}$ 的全息光栅作为系统的色散元件,输入窗口 $W_F=35\ \text{mm}$, 傅里叶频谱平面上的取样狭缝的宽度为 $70\ \mu\text{m}$, 400 A. S. A. 的 Kodak Ektachrome 彩色底片用于记录输出频谱。所得到的中文语音信号的频率假彩色频谱如图 6(a)、(b) 及 7(a)、(b) 所示(见彩色插页)。

普通的语音谱图是三维的,换句话说它显示了语音信号的时间、频率和振幅的信息。水平轴和垂直轴分别代表时间和频率的坐标,而谱线的明暗则是振幅坐标或强度的表示。上述结果具有明显的语音频谱图的特征。不同频率的谱线代表声带激励的不同谐音,它们随着声调的高低而呈现波状起伏。图 6(a)、(b) 是中文语音“光学信息处理”的谱图,但(a)出自女声,(b)出自男声。对比同一语音的男女声谱图可见,它们显示出具有类似的轮廓但稍不同的细节,男声谱图在低频范围内包含有更丰富的谐音成份。

上面的频率假彩色编码的谱图,低频显示蓝色,随着频率的增高由蓝到红而变化。改变系统中光栅及输入空间信号片相对于取样狭缝倾角 θ 的倾斜方位即可调节频率假彩色。谱图中泛白处则表示该频率在该时刻具有很强的能量。

由于所采用的阴极射线示波器的分辨率不高(约 $2\ \text{line/mm}$),因此空间信号片的高频响应极限只相当于 $2000\ \text{Hz}$ 左右,此外频谱图上频率坐标为非线性标度,这些均有待于进一步改进提高。

五、结束语

我们提出了用白光光学处理系统做语音信号的频率假彩色谱图的技术。这一系统明显地抑制了相干光学频谱分析系统的相干噪声,而频率假彩色编码的谱图更利于适应人的视觉系统对谱图的识别能力。如果能以声光调制器件代替空间信号片,以电视摄像系统取代记录底片,可以期望实现实时处理的白光光学频谱分析技术得到更广泛的应用。

作者感谢母国光教授对本文所提出的宝贵建议。

参 考 文 献

- [1] W. Koenig et al.; *The Journal of the Acoustical Society of America (J. A. S. A.)*, 1946, **18**, No. 1 (Jul), 19.
- [2] L. J. Cutrona et al.; *IRE Transactions on Information Theory*, 1960, June, 386.
- [3] C. E. Thomas; *Appl. Opt.*, 1966, **5**, No. 11 (Nov), 1782.
- [4] F. T. S. Yu; *IEEE Spectrum*, 1975, Feb., 51.
- [5] A. M. Tai et al.; *Proc. San Diego Biomed. Symp.*, 1977.

Speech spectrum analysis with a white-light optical processing technique

F. T. S. YU AND T. N. LIN

(Department of Electrical Engineering, The Pennsylvania State University, U. S. A.)

XU KEBIN

(Department of Physics, Harbin Institute of Technology, Harbin)

(Received 24 February 1984; revised 18 June 1984)

Abstract

The technique for generating color-coded speech spectrogram utilizing a white-light optical spectrum analyzer is described. First a conversion of the speech signal prerecorded on a magnetic tape to amplitude transmittance of photographic film is accomplished with a cathode ray tube (CRT) using density modulation method. The frequency pseudocolor-coded speech spectrogram is then obtained by means of a sampling narrow slit placed at Fourier plane in white-light optical spectrum analyzer. Operation of the system with theory of partially coherent light is discussed briefly in this paper. And some color-coded Chinese speech spectrograms are also provided.



第十六届国际高速摄影及光子学学术会议

第十六届国际高速摄影及光子学学术会议于1984年8月27~31日在法国马尔萨斯州的斯特拉斯堡(Strasbourg)举行。本届会议的东道国是法国,会议由法国国家技术研究协会(ANRT)负责组织,德国一些组织也积极参加了筹备工作,因此本届会议规定法文、德文和英文为工作语言。来自世界各地21个国家的450多名代表参加了会议,会议论文来自17个国家共207篇。中国参加会议的共9人(台湾省有4人),报送论文11篇(台湾省有1篇)。

会议宣读了九篇特邀报告,并且组织了十五个专题报告会。

从宣读的论文内容和数量可以看出,目前国际高速摄影及光子学领域中大致有下列几个特点:

1. 变像管微微秒技术仍然是高速摄影领域中最活跃、最受重视的部分。具体表现为:世界各先进工业国家均投入了相当的力量从事这方面的研究工作。最近连续四届获得沙丁(Shardin)奖的均为从事这方面工作的青年科学家。同时,在本届会议上将近1/4的论文是关于这方面的工作的。从论文的内容看,时间分辨率继续向亚微微秒方向提高;光谱响应往紫外、X射线及红外方向扩展,并且开始采用CCD代替SIT摄像管作实时读出。在整机方面开始考虑一机多用,以适应多种需要。

2. 高速视频技术方面论文的数量较上届有所增加,共17篇。除了介绍一些高速录像系统的性能外,大部分是在等离子体诊断、荧光图像分析、运动分析、X射线研究、医学、包装、建筑材料等工业方面的应用。也就是说由于它的即时放送、记录介质(磁带)的重复可用、记录时间长等优点,已在一些部门中开始与光机

(下转162页)