

音频监视



► 引言

有关图像的话题，我们可以谈论很多，但它只有和音频信号结合在一起，才能给人们提供一个完整的观看效果。如果人们在观看节目的同时，却发现音频有这样或那样的问题，这与有着相同的视频信号但无音频问题的节目相比较，人们肯定会认为，音频有问题的节目视频是很糟糕的。因此，我们必须给予音频信号以足够的重视，这是问题的实质。

音频器件既可以使用平衡信号，也可以使用不平衡信号。这两种信号都具有自身的物理特性和电特性，也有着各自的优势和不足。弄清这两种格式的音频信号，有助于我们掌握和应用音频信号。

不平衡信号

不平衡系统使用的是一根信号线和一根接地线。有时还采用屏蔽层。将不平衡的信号相互连接起来是很简单的，这只需使用价格相对比较低廉的电缆作导线，以及使用

如图1所示的RCA音频插座作连接器。连接器的外导体安装在音频设备上，通常与机壳相连。需要注意的是插头不要接反，以防止有源信号与机架地形成短路。不平衡音频信号的电路输入级也很简单，如图1所示。不平衡信号具有价格低廉和连接简单的特点，因此在家用产品中得到了广泛的应用。平衡信号具有共模抑制特性因而有着更低的噪声，但在家用产品的应用环境中是不必要的。不过，不平衡信号对外界干扰较为敏感，所以只能使用短电缆来传输信号。

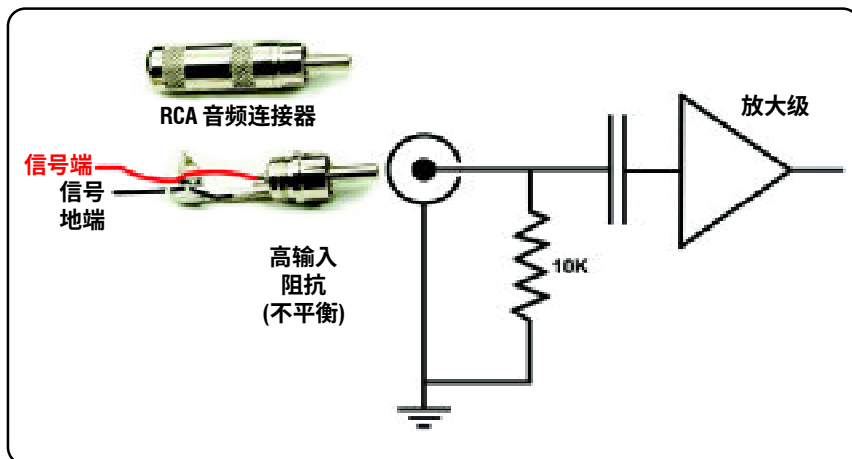
音频监视

► 应用注解

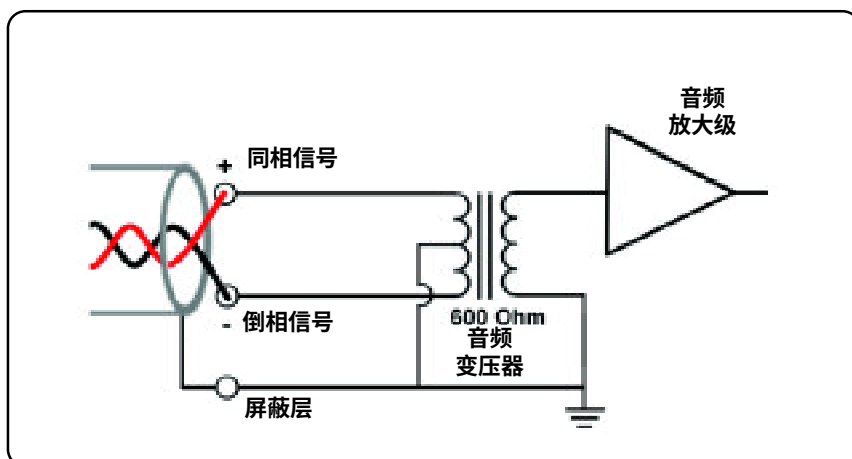
平衡信号

所谓平衡信号，指的是该信号具有两个分量，它们的幅度相等但极性相反，并且与连接器有着相匹配的阻抗特性。应用中通常将这两路信号分别称为同相信号和倒相信号，或正极性信号和负极性信号，或正向信号和反向信号。平衡信号之间的互连通常要使用三根导线，其中一对是传送同相信号和倒相信号的双绞线。将双绞线对用于信号传送，这样可使招致外界磁场干扰的环路面积最小。第三根导线就是屏蔽层。

屏蔽层可以衰减外界电场的干扰，尽管它不能完全消除外界电场的影响，但通过屏蔽层渗透而入的电场对同相信号和倒相信号产生的是相同的效果。这样渗透而入的信号就成为共模信号而加在电路的输入端。如图2所示。



► 图1. 不平衡音频信号间的连接



► 图2. 平衡音频信号间的连接

在平衡音频系统中，采用的是电子的或基于变压器的差分输入级放大电路。差分输入是从同相输入中减去倒相输入。正因为此，才使共模信号得以消除。平衡音频信号格式广泛应用于专业领域，由于它具有噪声抑制特性和高的信号幅度，使得它的相互连接较为复杂，相应也提高了成本。图3和图4中的XLR即为常用的典型连接器。

表1 平衡式输出结构配置

同相信号	XLR 插芯	2
倒相信号	XLR 插芯	3
屏蔽层	XLR 插芯	1

音频设备内部的电缆接线应当正确，以避免在设备内引起问题，如极性接反。

最简单的音频监视形式是使用能够显示音频信号幅度的电平表。有两种类型的电平表，即VU（Volume Unit）表和PPM表（Peak Program Meter），二者之间有着明显的差别。VU表和PPM对音频节目素材表现出不同的响应。VU表显示的是音频信号的平均音量电平，它具有对称的上升和降落时间，其积累时间相对较长（典型值为300ms）。积累时间主要由表内指针结构的机械惯性所决定。PPM表显示的是音频信号的峰值音量电平，它具有较快的上升时间（10ms）和较慢的降落时间（2.85s），其积累时间为10ms。PPM中的电子电路用以补偿机械摆动的惯性。由于存在着这些差别，因此VU表和PPM对音频节目素材一般有着不同的响应。

在用音频测试序列对系统进行调整时，在使用相同的音频节目的情况下，PPM应当比VU表有较低的读数才能使二者等效。广播管理部门发现，在PPM的峰值读数和VU表间有着良好的平均差值，该差值为8dB（分贝）。由此，

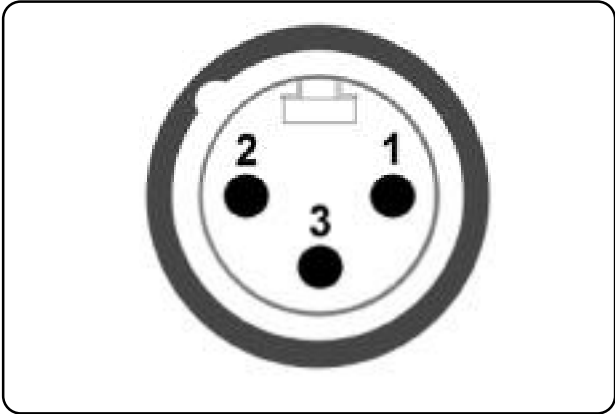


图3. XLR 插座。

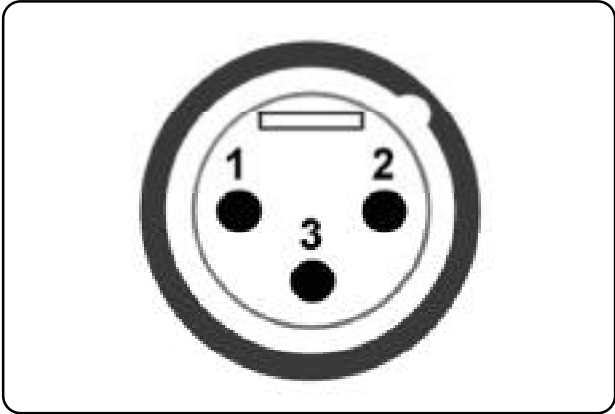


图4. XLR 插头。

他们规定，在进行音频调整时，VU表的0VU读数在PPM中应为-8dB。通过这样的调整，两种电平表对音频节目素材在实质上都有着相同的读数，但PPM能对节目峰值电平给出更为可靠的控制。

音频监视

► 应用注解

► 有关 dB 的换算

音频测量通常都用 dB (分贝) 来表示, 这是因为音频信号有着很宽的动态范围, 而 dB 可将电压或功率的测量值用对数的函数形式来表示。使用 dB, 还便于我们定量地改变音频信号幅度, 因为人耳的听觉与音频信号幅度呈对数关系。

$$\text{dB} = 20 \log \frac{V_2}{V_1} = 10 \log \frac{P_2}{P_1}$$

$$\text{注意: } P = \frac{V_2^2}{R}$$

V_1 = 参考电压电平

V_2 = 测量电压电平

P_1 = 参考功率电平

P_2 = 测量功率电平

P = 功率

V = 电压

R = 阻抗

在音频测量中也对 dBm 作了规定。dBm 是以 600 (负载上的 1mW 功率作为基准功率。这样, 在使用这个等式时, 0dBm 就相当于 600 (负载上的电压值为 0.775V。您可以计算出音频测量中常用的 dB 测量值, 它们用如下等式来表示:

$$\text{► dBm} = 10 \log P_1 / 0.001 \text{ W}$$

$$\text{► dBV} = 20 \log V_2 / 1 \text{ V RMS}$$

$$\text{► dBv} = 20 \log V_2 / 775 \text{ mV RMS}$$

$$\text{► dBu} = 20 \log V_2 / 775 \text{ mV RMS}$$

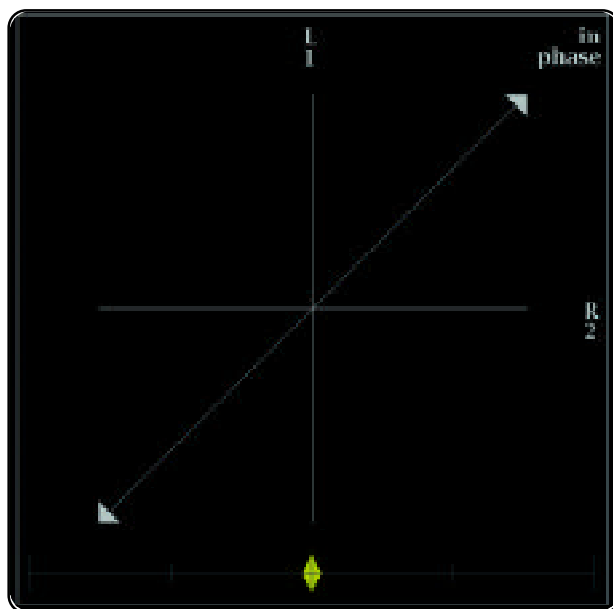
$$\text{► dB SPL} = 20 \log P_1 / P_2$$

设置音频电平

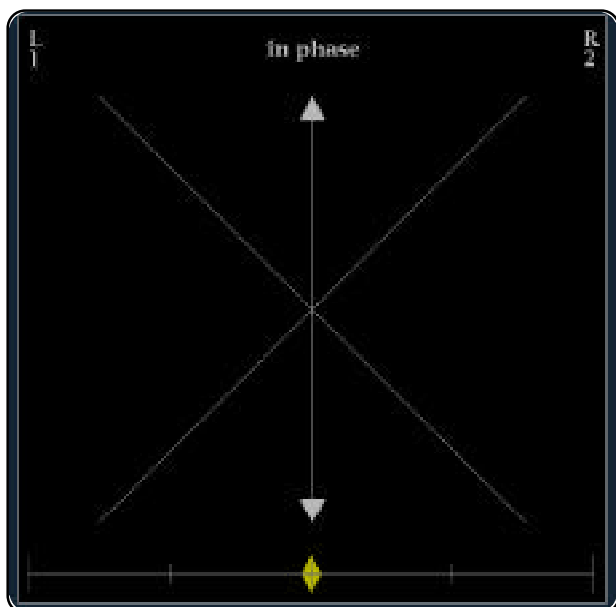
李沙如图形

将左右两声道的输入信号施加于 X-Y 显示器上, 其显示如同矢量仪。左声道加于 N-S 轴, 而右声道加于 E-W 轴, 如图 5 所示。许多音频专业人员对 “Sound Stage” (声平台) 模式更为熟悉, 它只需将上述图形旋转 45 度即可得到图 6 显示, 这样更便于观察两个声道的相位是否正确。李沙如图形对两个声道的再次混合的整个能量分布提供了即时的显示, 从李沙如图形中, 可以迅速地观察到这两个声道的混合能量分布是平衡式的还是集中偏向某一边。图 7、8 和图 9 描绘出不同的能量分布。

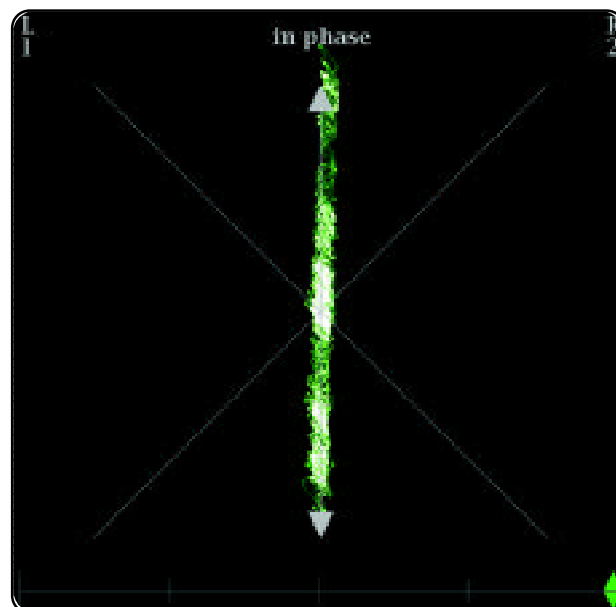
从音频监视器中可以清楚地发现音频节目素材中的问题, 例如音频信号中的削峰现象在李沙如图形中表现为矩形边缘的开口。图 10 表示的是音频信号经过严重削峰的情形。



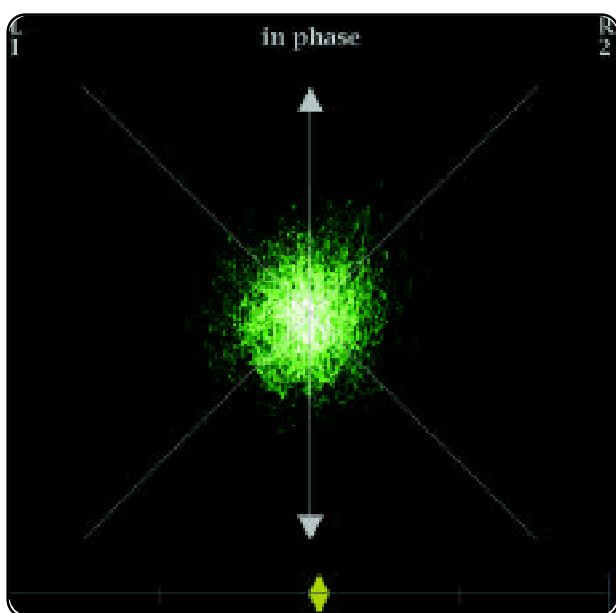
► 图 5. 李沙如图形的 X-Y 显示。



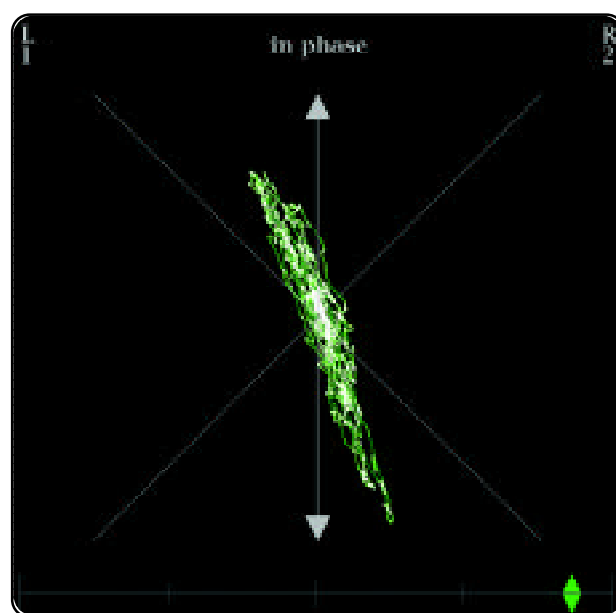
► 图6.“Sound Stage ”（声平台）模式中的李沙如图形显示。



► 图7. 合成器中的单音显示。



► 图8. 相关性较低的立体声信号。



► 图9. 左声道信号较强的立体声信号。

音频监视

► 应用注解

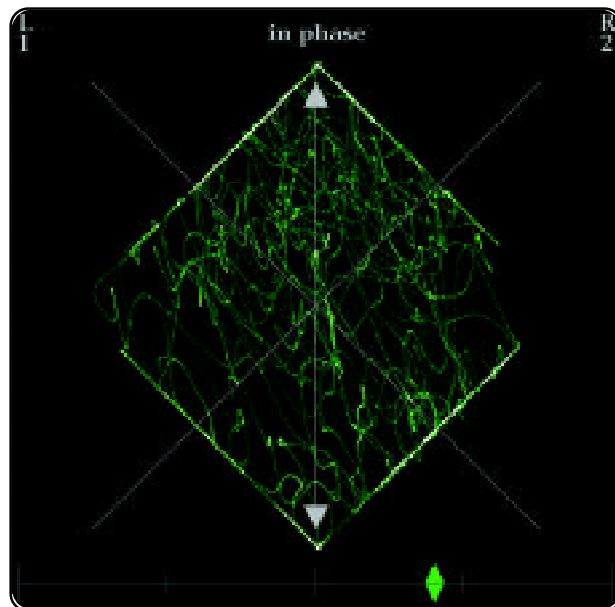
系统相位误差

系统间的相位误差会给音频信号带来许多不利的影响。使用音频监视器，可以帮助我们迅速地定量地判断系统相位间的误差。

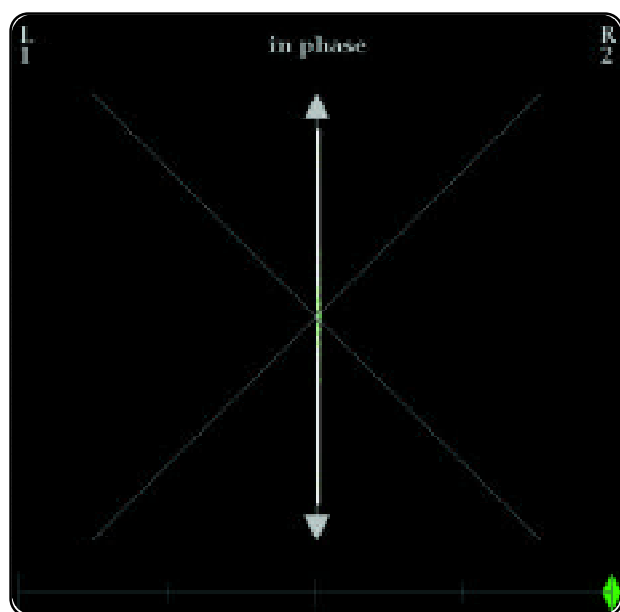
选择自动增益控制，使椭圆的边缘恰好触及相位切线。如果出现一条直线并与 $L=R$ 轴线相重合，那么被测设备的左右声道的音频信号相位相同、幅度相等，如图11所示。如果直线呈倾斜状，那么左右声道的相位相同但幅度不等。如果出现的直线与 $L=R$ 轴线相垂直，表示左右声道音频信号的相位相反。最后，如果椭圆的主轴落在 $L=R$ 轴线上，表示幅度相等但相位失配。

相反的信号极性

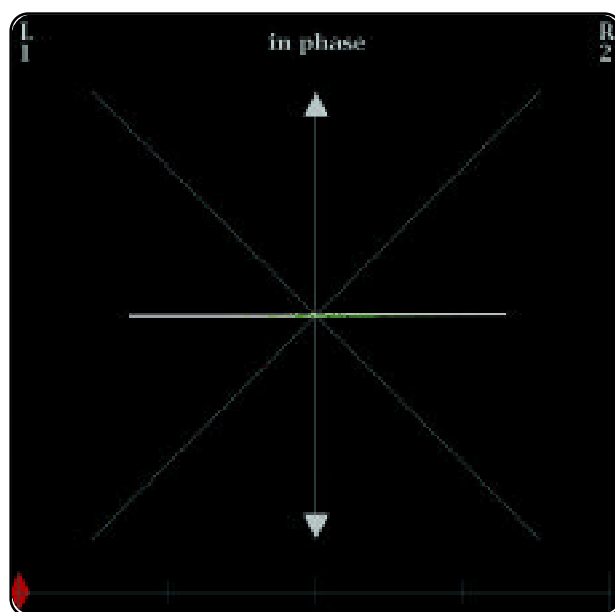
在较复杂的演播室中，多个话筒的使用将不可避免地带来相位极性不一致的问题，这样的情况是会经常发生的。不管什么时候，只要发生了极性相反的问题，就可使用音



► 图10. 经过严重削峰的立体声信号。



► 图11A. 左右声道的相位一致，增益相等。



► 图11B. 左右声道的增益相等，但相位相反。

频监视器快速跟踪直至信号源。将正弦波信号送入系统的各个通道，逐级对信号输出进行查验，可以迅速地查明极性相反的信号源。相位正确的信号在声平台李沙如图形上应当产生一条竖直线，如果相位相反，那么在李沙如图形上将会是一条水平线。

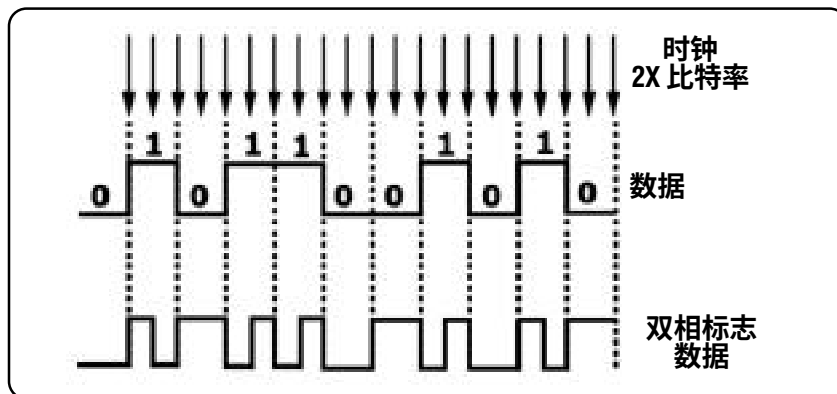
数字音频

按照视频业界的AES 3标准来传输数字音频已经延续多年。在 AES 3 标准中，其接口为串行数据码流，没有单独的时钟信号，为使接收机恢复数据，应当从传送的数据码流中抽取时钟信息。为此，只需采用一种简单的编码方案即可实现，这种编码称为双相标志编码，如图12所示。在双相标志编码中，每隔一个比特周期发生一次电平转换，当数据值为“1”时，则在后半个比特周期内再一次发生电平转换。采用双相标志编码，可以很容易地从数据中抽取时钟，并且使信号传输具有最少的直流分量。由于双相标志编码是用电平转换来表示数据值，因此这样的编码信号对极性也不敏感。AES 3标准支持多个取样频率，这些频率可以是32KHz、44.1 KHz(CD 用)和 48 KHz(专业用)，其中后者在视频设施中获得广泛应用。模拟音频信号按照时钟频率取样，并用16、20或24比特来表示音频信号的幅度。与视频信号相比较，音频信号所需要的量化比特数要大得多，这是因为音频信号有着更大的动态范围，同时增加量化比特数也可提供足够的信噪比(SNR)。

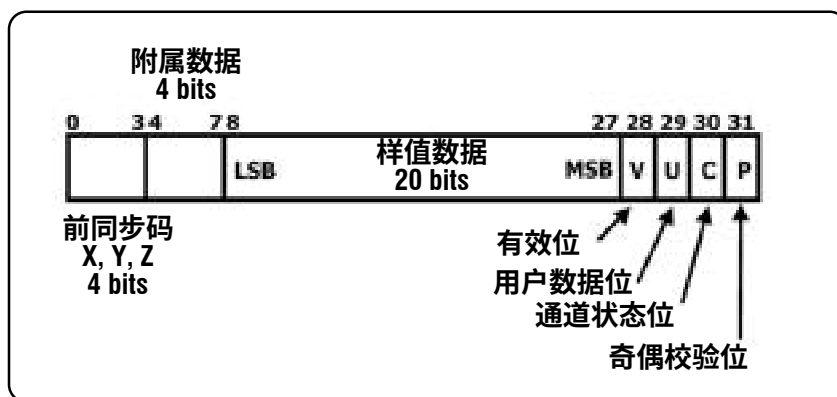
在数字音频中，决定 SNR 的基本公式为：

$$SNR = (6.02 \cdot n) + 1.76$$

此处“n”为每样值的量化比特数。



► 图12. 双相标志编码



► 图13. 子帧结构

对于16量化比特系统，理论上的最大SNR为 $(6.02 \times 16) + 1.76 = 98.08\text{dB}$ ，对于20比特器件，SNR为122.16dB，而对24比特器件，SNR为146.2dB。一个设计良好的20比特的模数转换器(ADC)，其给出的SNR典型值在100dB和110dB之间。

嵌入在串行数据码流中的数据包含两个音频通道，即通道1和通道2，并将它们复用在一起。这两个通道既可以是各自独立的单声通道，也可以是包含左右声道的立体声对。在单音频通道中，第二个通道的数据可以与第一个通道相同，也可以不包含数据而将数值设置为逻辑“0”。被称为前同步码的4比特同步字用于通道识别，它的编码方法与双相标志编码不同。这样做，有助于区分同步字与其它数据字。

音频监视

► 应用注解

每个通道的数据组合为 32 比特的子帧，如图 13 所示。通道 1 和通道 2 的这两个子帧经复用后形成一帧。将 192 帧进一步组合就形成一个块，图 14 即为帧格式。

有关子帧的组成可描述如下：

前同步码

前同步码是 4 比特的同步字，它用于通道识别并作为音频数据块的开始。通道 1 以前同步码 X 作为标识，通道 2 以前同步码 Y 作为标识。至于含有 192 帧的数据块，则以前同步码 Z 作为标识，如图 15 所示。前同步码未采用音频数据的双相标志编码，这样便于将它与其它数据字区分开来。

表 2 给出了用先前“0”或“1”态表示的前同步码数值。通常情况下，在上述数值中，只能传输一种类型的数值。然而，信号通道中的极性反转需要对这两种状态进行解码。

表 2 给出了用先前“0”或“1”态表示的前同步码数值。通常情况下，在上述数值中，只能传输一种类型的数值。然而，信号通道中的极性反转需要对这两种状态进行解码。

表 2 前同步码的定义

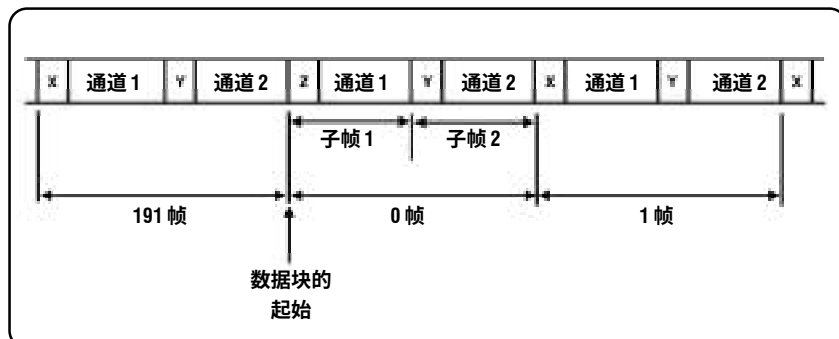
	先前状态 0	先前状态 1	标识
X	11100010	00011101	子帧 1
Y	11100100	00011011	子帧 2
Z	11101000	00010111	子帧 1 和块起始

附属数据 (Auxillary Data) 比特

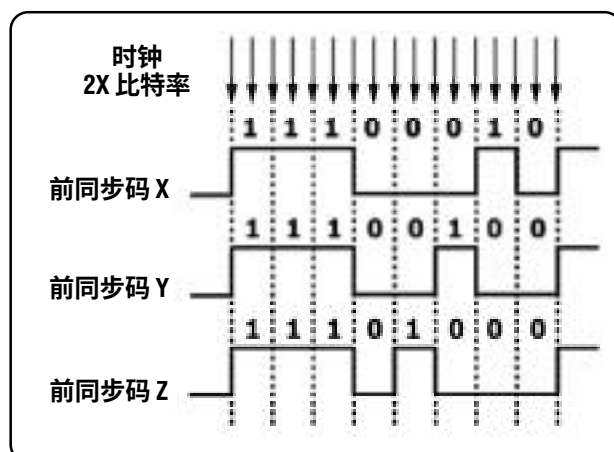
如果只使用 20 比特的音频样值，则前面的 4 个最低有效位 (LSB) 则作为附属数据使用。这些附属数据比特的一项应用是用作语音质量 (voice-quality) 音频通道，以提供内部通话通道。否则，可将这 4 个比特作为 24 比特音频样值的四个最低有效位。

音频样值数据比特

音频样值数据位于比特 4 和比特 27 之间，其中最高有效位 (MSB) 位于比特 27，音频样值数据最多可支持 24 比特的样值。如果在这 24 比特未全部用作音频数据样值，那么



► 图 14. AES/EBU 的帧格式



► 图 15. 前同步码 X、Y 和 Z。

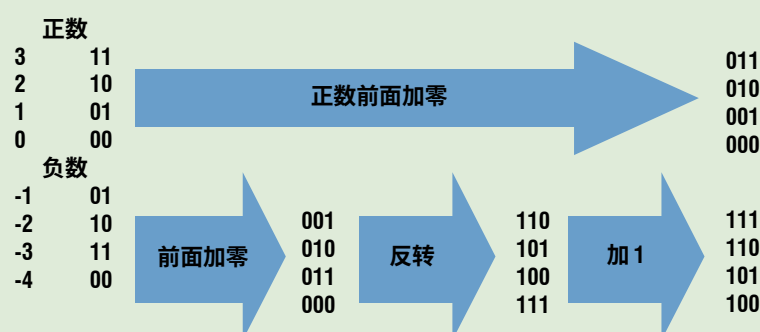
LSB 的 4 个比特设置为“0”。在广播设施应用中，一般只使用 20 比特的样值。这样，由比特 4 至比特 8 的 4 个最低有效位就可用于附属数据通道。

20 比特的音频样值在广播环境中得到了最广泛的应用。而 24 比特音频样值也被 AES/EBU 所支持，这样从比特 4 至比特 27 全部用于音频数据样值而不提供附属数据比特。二进制音频数据样值是二进制补码。使用这种简单的技术可大大降低音频硬件设计的复杂性。

子帧也传送附加数据比特，它们可提供音频通道的有用信息。

► 二进制补码

在二进制补码中，使用最高有效位（MSB）以表示信号的正负属性。如果MSB 设置为“0”，那么表示该数值为正数；如果MSB 设置为“1”，那么表示该数值为负数。在下面的例子中，给出了二进制补码中二位数的可能范围。



有效标志位 (V)

如果将有效标志位设置为“0”，则子帧音频数据适于解码为模拟音频。如果将有效标志位设置为“1”，则音频样值数据不适于解码为模拟音频信号。可将测试设备设置为忽略有效标志位，这样就可以继续使用数据而进行测量。

用户数据位 (U)

用户数据位可用于传送音频信号的附加信息。在192个子帧中，每一个用户数据位可以组合在一起，以在每个音频数据块中产生总数为192位的数据。使用者可以利用这些信息，例如可用作版权信息。

通道状态位 (C)

通道状态位提供与音频信号相关的各种参量。在每个音频通道中，将192个子帧中的每个通道状态位组合在一起以用于这些参量。表 3 给出了这些比特所传送的信息。

通道状态位具有三种级别：最小级、标准级和增强级。建议将标准级作为专业电视应用，因为标准级的通道状态位包含有关信号加重、取样频率、通道模式（立体声、单声等）、辅助比特的使用（将音频数据扩展为24位或作其它应用）的信息，此外还含有用于整个通道状态块错误校验的CRC（循环冗余码）。有关通道信息细节描述的附加信息，可参见附录 A。

奇偶校验位 (P)

奇偶校验位的设置是将子帧中比特 4 至比特 31 的数值形成一个偶校验(偶数个1)，它是校验错误的一种简单方法，用以检测子帧中的一个错误。注意对有前同步码的子帧来说并非必需，因为前同步码中已经具有偶校验。

音频监视

► 应用注解

表3 通道状态应用

比特字节	0	1	2	3	4	5	6	7
0	通道状态 通道的应用	音频 / 非 音频应用	音频信号增强			源取样频率 的锁定	取样频率	
1	通道模式				用户比特管理			
2	辅助样值比特的使用			源字长度和源编码历史			保留	
3	未来多通道功能描述							
4	数字音频基准信号		保留					
5	保留							
6	字母数字通道原始数据							
7								
8								
9								
10	字母数字通道目标数据							
11								
12								
13								
14	本地样值地址编码							
15								
16								
17								
18	时间 / 日期样值地址编码							
19								
20								
21								
22	可靠性标识							
23	循环冗余校验特性							

AES 互连

在传送AES/EBU串行数字数据时，可以采用两种基本的连接方式。一种连接方式是使用标准的XLR连接器和双绞线电缆来传送数字信号。在XLR连接器中，2芯和3芯传送平衡数据信号，而1芯则用作屏蔽。注意由于被传送的信号对极性不敏感，因此双绞线中哪一根与2芯或3芯相接无关紧要。然而按照惯例定义2芯接“正极”而3芯接“负极”。

表 4. 平衡输出配置

信号“+”	XLR 芯	2
信号“-”	XLR 芯	3
屏蔽	XLR 芯	1



► 图16. XLR 插头连接器

图 17 为一标准的互连电路。其输出电压在 2 和 7V 之间，阻抗为 100W，信号的传送采用了变压器耦合并为平衡式。在这样的电路中，数字音频信号可以通过 100 米电缆来传送而无需均衡器。

另一种连接方式是按照 AES3 — ID 标准的定义，使用标准的 75W 同轴电缆和 BNC 连接器。这是一种不平衡的接口，它可以在标准的同轴电缆上传输 AES/EBU 数字音频信号。这种连接方式可以利用现有的广播基础设施。

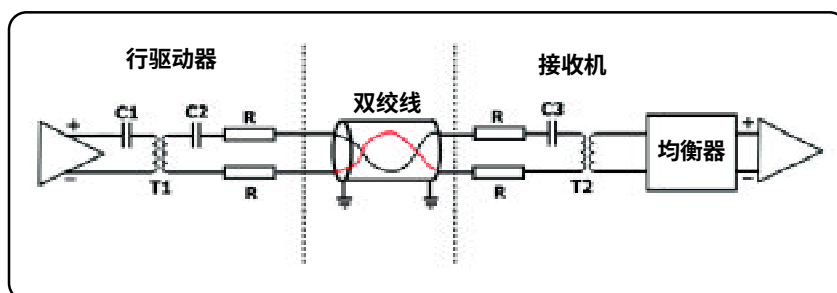
在正常情况下，信号需经变压器耦合。但这对于不平衡信号却并非必需。图 18 为一标准的互连电路。电路的输出电压为 1.0Vp-p，阻抗为 75W 不平衡式。在这样的电路中，信号可以通过 1000 米的电缆来传输。为了在 XLR 和 BNC 接口之间实现互连，可以采用各种形式的电路，既可使用简单的电阻匹配网络，也可使用变压器和衰减器电路。

设置电平

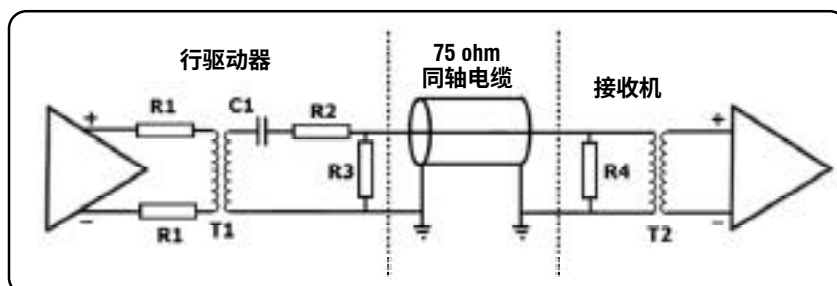
在设置电平和对数字音频信号进行说明时，要弄清几个问题。在音频信号取样的过程中，数字音频的最大值是用全“1”的样值来代表的，该值即所谓 0dBFS（分贝满刻度）。如果原始的模拟音频信号超出此值，就有可能出现削峰，并在音频信号中产生失真。还有，当数字音频又还原为模拟信号时，数字音频信号也有可能产生较高的模拟幅度。这是因为在转换的过程中，模拟输出级后的低通滤波器会给出较高的模拟幅度信号电平，它要高于数字数值所代表的信号电平。

因此，大多数音频监视器中提供一个内插的音频信号图示，用以指示信号峰值，并给出削峰发生时的指示。

除了 VU 表和 PPM 表之外，模拟音频监视器上的冲击式电平表和数字音频监视器提供了真实的峰值指示。这种类型的显示给出了几乎是瞬时的实际信号峰值，而不管间隔时间的长短。在冲击式电平表中，用户可以选择适当的参考电平和峰值电平以满足他的特定要求。一个实际例子是，将 -18dBFS 作为测试电平，将 -8dBFS 设置为峰值节目电平。用黄色的菱形表示测试电平，用红色菱形表示峰值节目电平，这些菱形均分布在条形电平指示的旁边。这样，操作者很容易地对电平作出说明。当音频信号在测试电平参考值以下时，电平呈绿色条指示，而黄色条指示信号电平在测试参考值以上，红色条指示信号在峰值电平以上。图 19 是带有数字音频选项 DG 的 WFM700 的音频典型显示。



► 图 17. AES3 互连电路。



► 图 18. AES3 — ID 互连电路。

音频监视

► 应用注解

上述显示还可给出削峰和静音指示以及可由用户编程的两个门限。

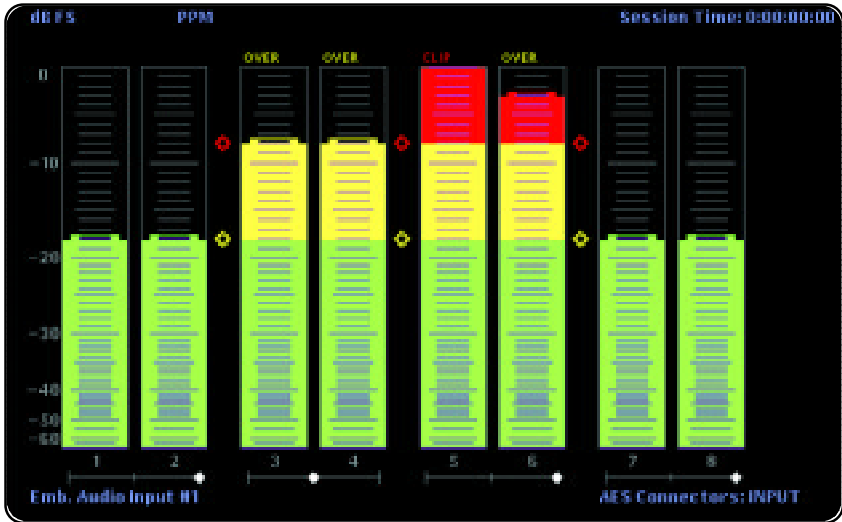
削峰 (Clips)、静音 (Mutes)、电平过高 (Over)

在数字音频中，多个连续样值的音频取样数据为 0dBFS 时，就会出现削峰状态。当削峰出现时，在用户设定的时间周期内，电平显示器将给出状态指示。

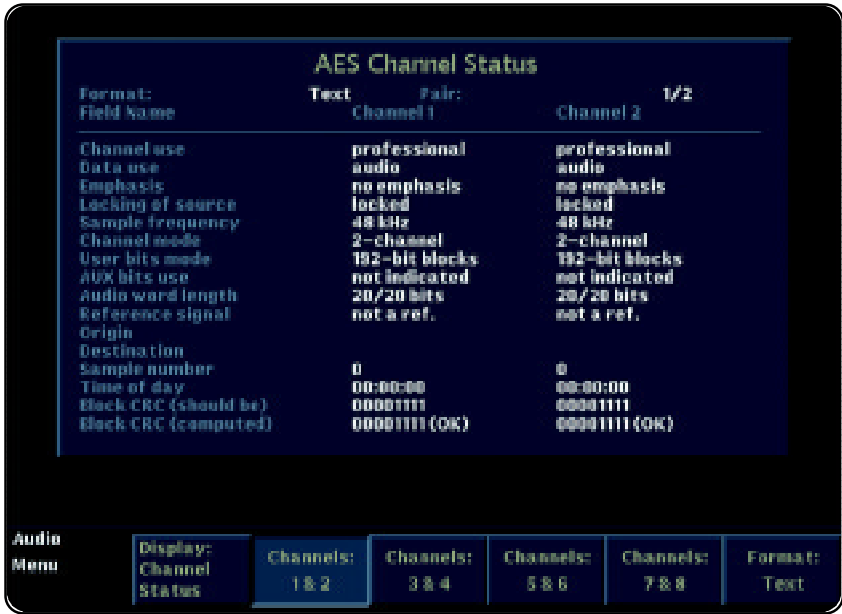
静音状态是指通道中如果有多个连续样值的数字音频数据保持为零值时，在条形显示器中所给出的状态指示。电平过高是指当音频信号超出用户设置的电平值时所出现的状态。同样地，如果音频信号低于用户设置的电平值时，即会出现寂静 (silence) 状态。当被监测设备出现上述各种情形时，数字音频监视器就会作出记录并显示在屏幕窗口中，同时给出用户监视期间内的所发生的错误类型和错误次数。

数字音频中的另一类型的失效模式是音频样值有可能受到损坏。如果出现这种情况，那么该音频通道子帧中所含的有效标志位就会设置为“1”，用以指示错误的发生。同时，音频设备将忽略遭受破坏的数据并使相应样值的音频信号保持为静音。如果用户希望观察到所有的音频数据样值，那么可在音频监视器中加以设置以忽略此有效位，这样就可以使用信号中的音频数据样值而进行测量。

可以用各种方式来显示通道状态。图20就是一个例子，它是 WFM700 (含选项 DG) 所给出的通道状态的文本描述。并按照 AES/EBU 3 规范的定义作出有关数据的解释。



► 图19. 音频条电平。



► 图20. AES 通道状态的文本格式描述。

还可对通道状态显示作出配置，以用二进制形式 (图21)、十六进制形式或XMSN二进制形式对各个特定数据值逐个查看。

SD/HD 中的嵌入音频

在串行数字视频信号的辅助数据 (Ancilliary data) 区内,可以嵌入AES/EBU 音频数据。这对于一些大型系统是很有用的。在这些大型系统中,需要考虑配置单独音频路由的造价问题,而且要确保音频信号与其相对应的视频信号间有着良好的相关性。至于小型系统,例如节目后期制作的配套设施中,维持单独的音频路由要更经济些,这样可以不必使用大量的嵌入器和解嵌入器模块。我们仍将集中讨论串行数字分量信号,尽管有关标准定义的是数字复合信号中的数字音频嵌入方法,这一课题在此不作讨论。有关数字复合信号中的数字音频嵌入信息,请参看 *A Guide to Digital Television Systems and Measurements* (数字电视系统和测量指南,文献号为 25W-7203-3)。

在标准清晰度 (SD) 和高清晰度 (HD) 视频格式中,串行数字接口 (SDI) 中的音频数据复用由以下标准所定义:

- **SMPTE 272M – 数字视频辅助数据区中的AES/EBU 音频和辅助数据格式**,适用于按照 SMPTE 259M 定义的标准清晰度格式
- **SMPTE 299M – HDTV Bit 串行接口中的24-Bit数字音频格式**,适用于按照 SMPTE 292M 定义的高清晰度格式



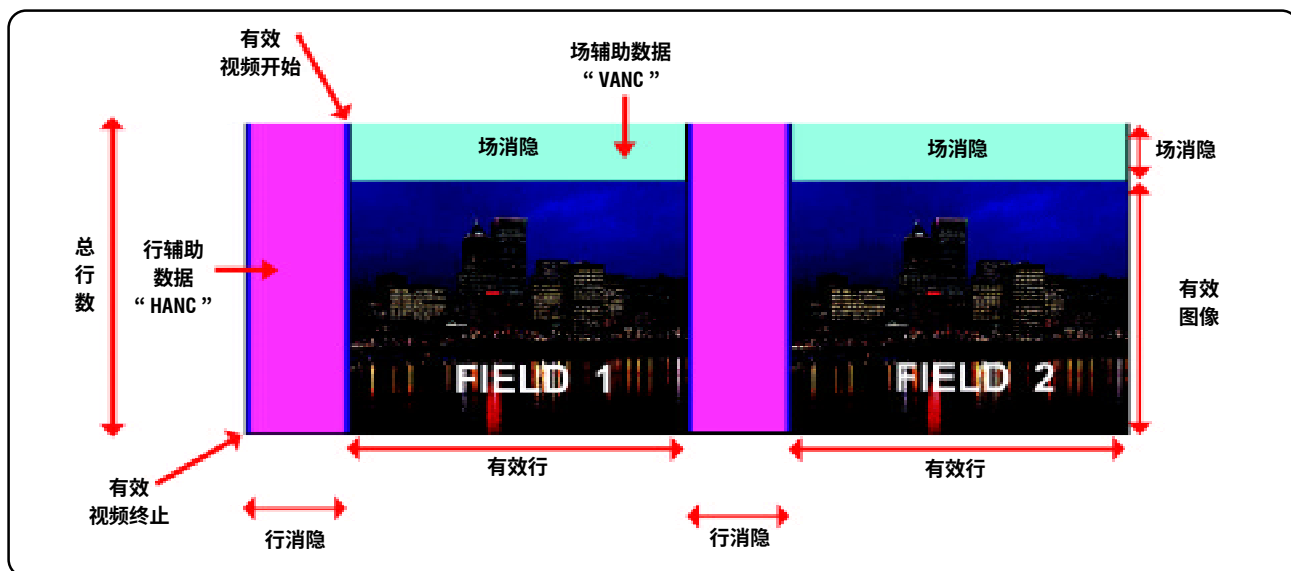
► 图21. AES 通道状态的二进制格式。

辅助数据 (Ancilliary data) 区

用于分量数字视频中的辅助数据区如图22所示。在行、场消隐区中,除了少量用于EAV (有效视频终止) 和 SAV (有效视频起始) 序列以及HD中用于行序号和CRC信息的数据之外,其余均可用于辅助数据。可将辅助数据区划分为两类 — HANC (行辅助数据) 和 VANC (场辅助数据)。在标准清晰度格式中,音频数据划分为20比特音频样值,也可将4个附属数据比特扩展为音频数据。而在高清晰度格

音频监视

► 应用注解



► 图22. 分量数字格式中的辅助数据区。

式中, 将所有的24比特音频数据用一个数据包来传送。对于标准清晰度格式和高清晰度格式, 音频数据均位于行辅助数据区内。对于SD系统, 附加的扩展数据也在HANC中传送。在SD中, 使用Cb/Y/Cr/Y' (Y是共有的亮度样值, 而Y'是独立的亮度样值) 来传送音频数据, 而在HD格式中, 仅使用Cb/Cr数据字传送音频数据, 而Y样值则用于传送可选的音频控制包, 音频控制包位于切换点之后的第二行, 每场发送一次。如果使用48KHz时钟的音频数据速率, 并不一定需要音频控制包。在相应视频格式的切换点期间, HANC中没有音频数据。在标准清晰度格式中, 在错误检测和处理插入期间, 也没有音频数据。

在HANC中, 按规定最多可以嵌入16个通道的音频数据。这16个通道分为4组, 每组含有4个音频数据通道。

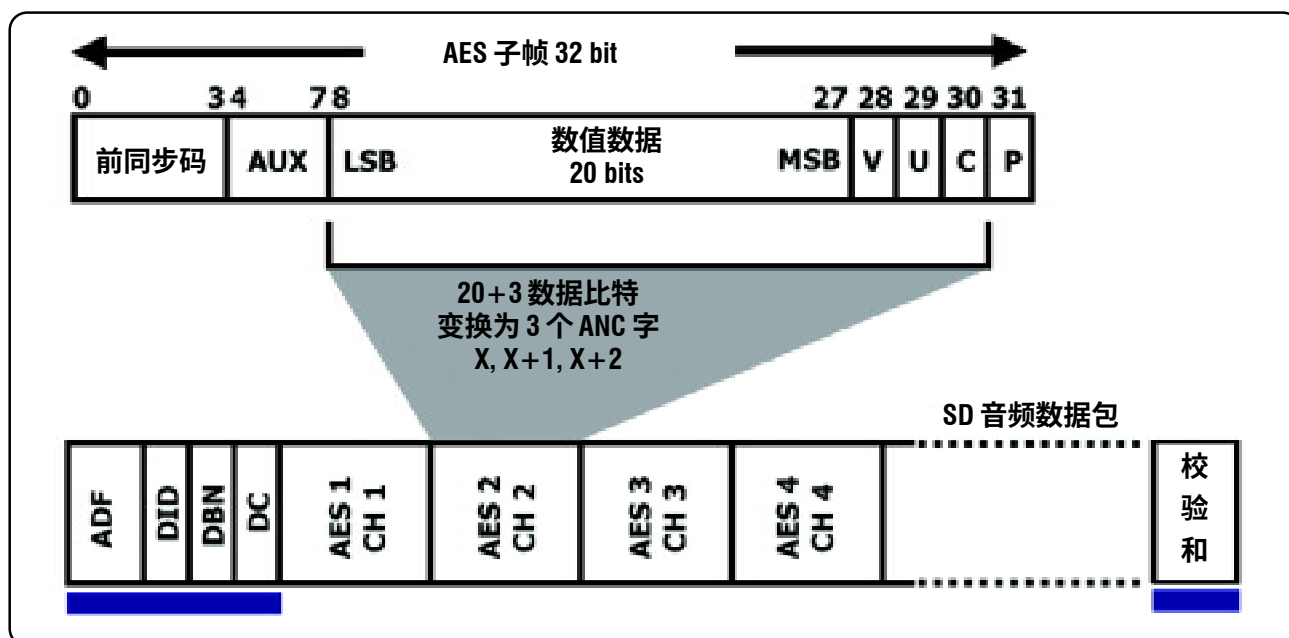
始处是使用数字值的包头, 它不含数字视频数据而作为同步保留字。包头中的辅助数据标志 (ADF) 含有三个字: 000_h 、 $3FF_h$ 、 $3FF_h$ 。通过不同的数据识别符 (DID, 一个字) 来区分各种类型的数据包。数个不同的DID字经定义以组成用于嵌入音频的各个数据包。数据块号 (DBN, 一个字) 是可选用的计数器, 可给辅助数据包提供顺序指令, 使接收机确定数据是否丢失。例如, 在嵌入音频中, DBN序列可用来检测是否发生了场消隐切换, 使接收机对音频数据进行处理以除去瞬间的类似“卡塔”声或“爆裂”声。数据包中表示数据量的数据计数字 (DC) 恰在用户数据之前。最后, 用户数据之后是一个字的校验和, 它用来检测数据包中的差错。

辅助数据的格式化

辅助数据先经格式化形成数据包, 然后再经复用进入视频数据码流, 如图23所示。每个数据块最多可含有255个用户数据字, 多个数据包可以分别放置在各个辅助数据区内, 这样便可提供相当灵活的数据通讯通道。在每一数据包的起



► 图23. 辅助数据的格式。



► 图24. 基本嵌入音频的格式。

基本 SD 嵌入音频

嵌入音频由 SMPTE 272M 所定义，它可提供多至 16 个通道、20 比特的音频数据，取样频率为 48KHz，取样时钟锁定于电视信号。尽管这些规定是在该标准的复合数字部分作出的，但同样适用于分量数字视频。这种基本的嵌入音频对应于嵌入音频标准中的 A 级。其它的运行级别可提供更多的通道，具有不同的取样频率以及有关音频数据的附加信息。基本嵌入音频数据包的格式源自于 AES 音频，如图 24 所示。

音频数据包包含有四个音频通道的一个或多个音频样值。AES 的每个子帧中的 23 个比特（20 个音频比特再加上 C、U、V 比特）变换为三个 10 比特的视频字（X、X+1、X+2），如表 5 所示。其中比特 9 总是比特 8 的逻辑非，以保证已除去的字数值（3FF_h 至 3FC_h 或 003_h 至 000_h）不被使用，Z-bit 预置为“1”，它对应于 192 帧 AES 块的第一帧。各个嵌入音频的通道基本上是独立的（尽管它们总是成对传送），因此把每个通道中 Z-bit 预置为“1”，即使它们均来自于同一 AES 源。C、U 和 V bit 是由 AES 信号变换而来的，不过，它们的奇偶检验位（P*）不是 AES 的奇偶检验位（P）。数据字 X+2 中的 Bit-8（P*）是所有三个字中 0-8 比特的偶校验。有关音频数据包的分配现有一些规定和限制，尽管

在标准的早期条款中考虑到一些原有设备可能并不满足所有这些规定。按照 RP168 的规定，在正常场消隐期切换后的行辅助数据区内，是不传送音频数据包的。RP165 规定，在设计用于误码检测检验字的辅助数据区内，也不传送音频数据包。考虑到这些限制，除去上述不传送音频数据包的数据区，要求“音频数据在整个视频场内应当尽可能地平均分配”。对于基本级别 A，作上述处理后，在每个音频数据包内，每个通道就有三个或四个音频样值。

表 5. 嵌入音频的比特分布

比特	X	X + 1	X + 2
B9	B8	B8	B8
B8	音频 5	音频 14	P*
B7	音频 4	音频 13	C
B6	音频 3	音频 12	U
B5	音频 2	音频 11	V
B4	音频 1	音频 10	音频 19(MSB)
B3	音频 0	音频 9	音频 18
B2	通道 1	音频 8	音频 17
B1	通道 2	音频 7	音频 16
B0	Z- 比特	音频 6	音频 15

音频监视

► 应用注解

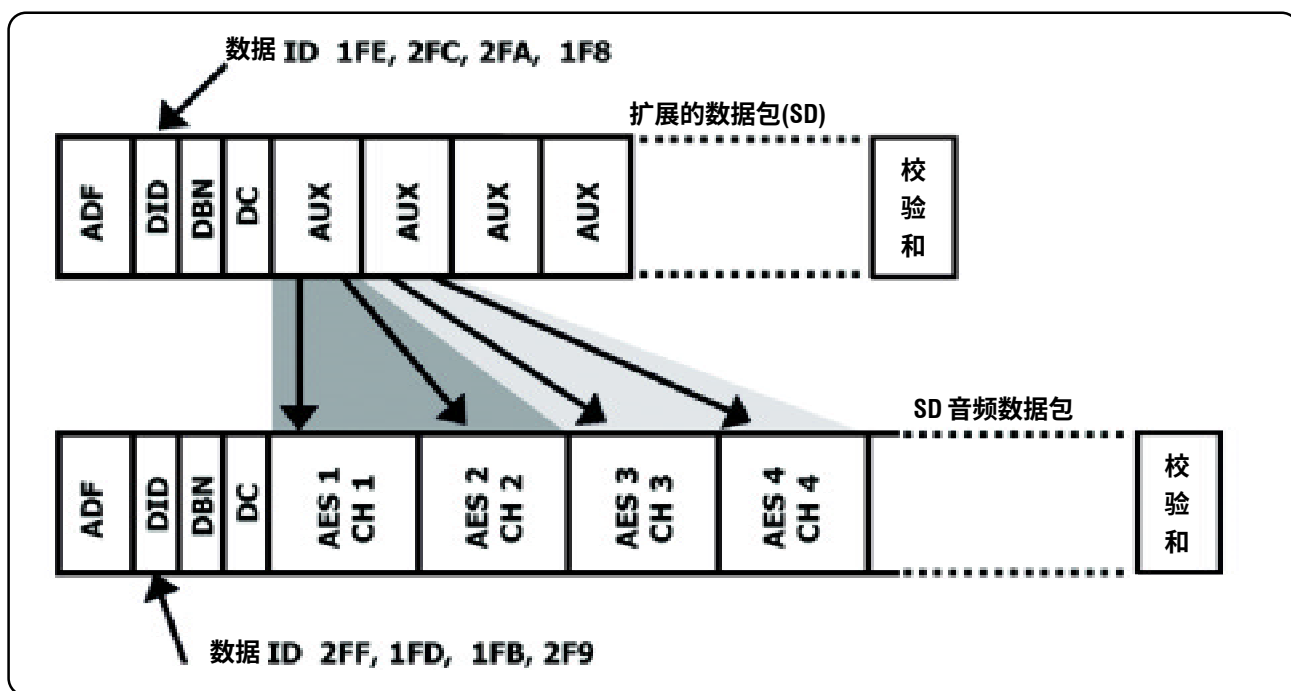
扩展的嵌入音频

扩展的嵌入音频具有如下特点：

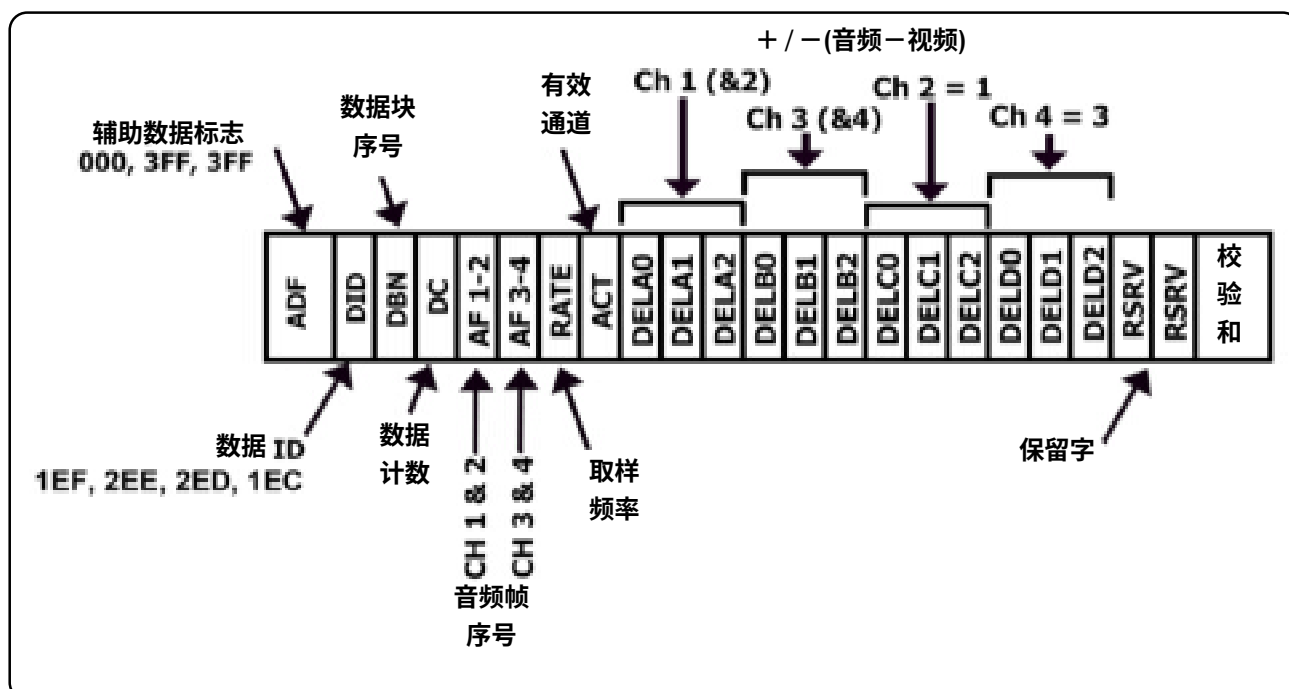
- 携带 4 个 AES 的附属比特(它们可将音频样值扩展至 24 比特)
- 允许运行于非同步时钟状态
- 取样频率可不同于 48KHz
- 为每个通道提供音频 / 视频的延时信息
- 编制数据识别符(ID), 允许分量数字系统中的音频通道数多至 16 个
- 525 行系统的“音频帧”计数

为了提供这些特点, 需要定义两个附加数据包。扩展的数据包携带 4 个 AES 附属比特, 采用的格式应当是这样的, 即扩展数据包中的一个视频字含有两个音频样值的附属数据, 如图 25 所示。扩展的数据包应当处于与其相关的音频数据包的相同辅助区内, 且应在音频数据包之后。

音频控制数据包如图 26 所示, 每场传送一次, 它位于场消隐切换点之后的第二个行辅助数据区内, 音频控制数据包所含信息包括: 音频帧号、取样频率、有效通道和各通道音频/视频间的相对时延。对于 48kHz 同步运行模式, 音频控制包的传送是可选的, 但对其它所有的运行模式(因为它包含有正在使用模式的有关信息), 则是必需的。在 525 行、29.97 帧/秒运行系统中, 音频帧号是人为规定的。在这样的系统中, 准确的 5 帧中含有准确的 8008 个音频样值, 这意味着每帧的样值数不是整数。在 PAL 制 625 行系统中, 其格式是 25 帧/秒, 这一过程就十分简单: 每帧中含有准确的 1920 个音频样值, 这是因为它的帧频是偶数。在音频样值数是整数的情况下, 音频帧的序列号为帧数。在 525/29.97Hz 系统中, 这种情况为 5 帧, 音频帧号指示某一特定帧在该序列中的所处位置。而在 625/50Hz 系统中, 这种情况为 1 帧。对于信号源的切换而言, 帧号是十分重要的, 因为某些设备, 例如数字 VTR, 需要稳定一致的同步运行, 以避免缓冲器的上溢和下溢。如果是用



► 图25. 扩展的嵌入音频



► 图26. 音频控制包格式

于频繁的切换，那么在设计接收设备时，应在音频帧序列被破坏处为五分之四的情况下，紧随切换之后增加或减少一个样值。关键是在这样的系统中检测已经发生的切换。使用辅助数据格式结构中的数据块号，以及在音频控制包的音频帧号字中，通过未用比特的可选帧计数器，将有利于问题的解决。音频控制数据包中包含有音频延时信息，它使用的是默认通道对模式。其中时延-A（DELA0-2）用于通道1和通道2，除非通道2的时延与通道1并不相等。在这种情况下，通道2的时延位于时延-C。在一个通道对中，每个通道的抽样频率应当相等，因为“ACT”中的数据仅提供两个数值，其中一个数值用于通道1和2，另一个数值则用于通道3和4。在分量数字系统中，为了提供多至16个音频通道，可将嵌入音频划分为与基本四通道运行相对应的音频组。在三种数据包（音频数据包、扩展的数据包和音频控制包）类型中，每一种类型均设计为四个数据ID，如表6所示。

表 6. 16 通道运行模式中的数据识别字

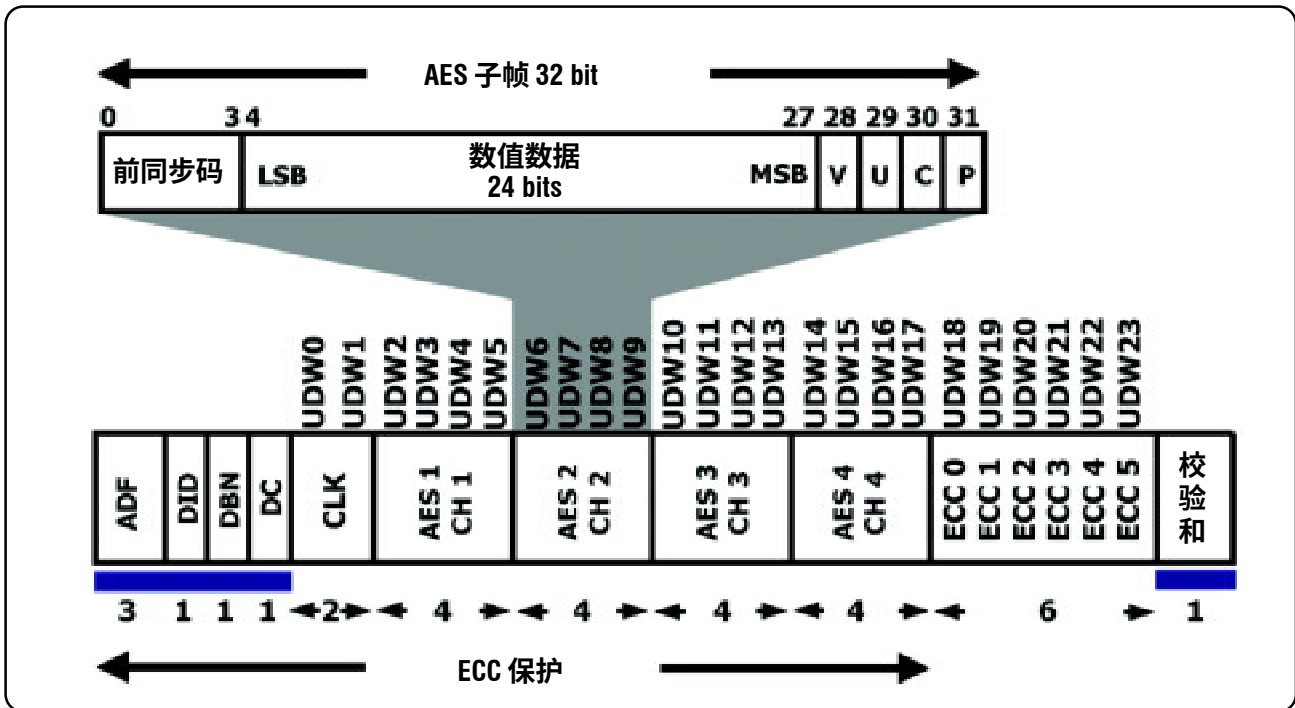
	音频通道 数据包	音频 数据包	扩展的 控制包	音频
第1组	1-4	2FF _h	1FE _h	1EF _h
第2组	5-8	1FD _h	2FC _h	2EE _h
第3组	9-12	1FB _h	2FA _h	2ED _h
第4组	13-16	2F9 _h	1F8 _h	1EC _h

接收机缓存器的大小

在分量数字视频中，音频复用器中的接收机缓冲器并非是一个很严重的问题，因为有足够的辅助数据区可供使用。为此在标准中规定需要每通道64个样值的接收机缓冲器，而早期的规定是每通道48个样值，以此提示设计人员在原有设备中存在一些限制。在标准中，级别A规定样值的分配允许使用每通道48个样值的接收机缓存器，而在其它级别中，一般需要使用已规定的64样值缓存器。

音频监视

► 应用注解



► 图27. HD 音频数据包的结构。

基本 HD 嵌入音频

和标准清晰度格式中的嵌入音频相比较，HD 环境中的 AES/EBU 应用既有相似之处，也有一些差别。其中，辅助数据包的格式在 SD 和 HD 之间是一样的，但数据包中用户数据所包含的信息却不相同。在 HD 中，整个 24 比特的音频数据是作为一组发送的，而不再划分为 20 比特的音频数据和一个含有 4 个附属比特的扩展数据包。因此在 HD 中所使用的比特总数为 29 比特 (SD 为 23 比特)，其中的 24 比特音频数据和 C、V、U 以及 Z-bit 标志一起放置在四个辅助数据字中。此外，HD 音频数据包中还包含有 CLK (Audio Clock Phase Data, 音频时钟相位数据) 和 ECC (Error Correction Code, 错误校正码)，如图 27 所示。由于所有的 24 比特音频数据是在用户数据中传送的，因此在 HD 中不再使用扩展的数据包。

HD 的辅助数据包结构符合 SMPTE 291M，它的辅助数据标志 (ADF) 为三个字，其数值分别为 000_h 、 $3FF_h$ 和 $3FF_h$ 。DID

为一个字，其数值如表 7 所示，它用于区分对应的音频数据组。DBN 为一个字的数据计数，其值总为 218_h 。用户数据字 (UDW) 部分总是包含 24 个数据字 (UDW0 至 UDW23)，其结构如图 27 所示。其中的前两个数据字即 UDW0 和 UDW1 用于音频时钟相位数据，并提供产生音频取样时钟的一种方法。这两个字中的数据用于计算 EAV 的第一个字和视频样值 (这个视频样值与该音频样值相对应) 之间的视频时钟数。其它 16 个数据字 (UDW2 至 UDW17) 用于携带 AES 音频数据，每四个 UDW 用户字中的各音频数据子帧的比特分配如表 8 所示。

表 7. HD 嵌入音频的 16 通道运行模式中的数据识别字

	音频通道	音频数据包	音频控制包
第 1 组	1-4	$2E7_h$	$1E3_h$
第 2 组	5-8	$1EG_h$	$2E2_h$
第 3 组	9-12	$1E5_h$	$2E1_h$
第 4 组	13-16	$2E4_h$	$1E0_h$

注意，由表8可见，并非所有的前同步码是在这四个字中传送的，而只有用于192帧的起始标识即Z-bit指示位是在这四个字中被传送。另外，在32比特子帧中所使用的奇偶检验位，也和标准定义有所不同。

ECC是一组6个字的序列，它用于检测前面24个字（由ADF至UDW17）中的错误。其值是这样计算出来的：使用这24个字中的数据B0至B7的8个比特，通过BCH编码（一种错误校正方法）信息传送线路，从而产生6个字的错误校正码。

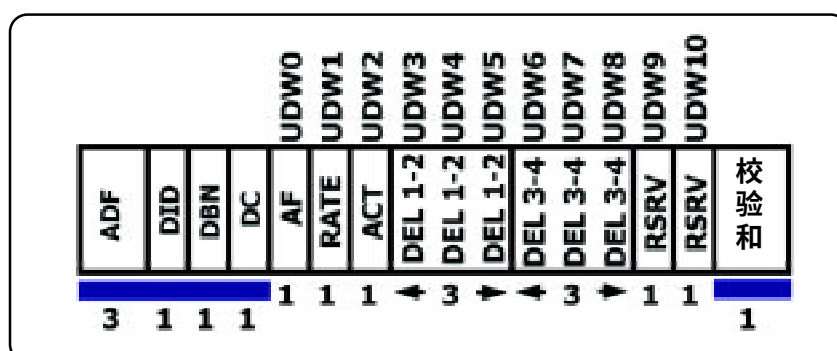
辅助数据信息仅在色差信号Cb/Cr数据区被复用，这点与标准清晰度格式的结构有所不同，在标准清晰度格式中，采用Cb/Y/Cr/Y'数据字传送辅助音频数据。在HD中，Y数据区仅用于传送音频控制包。音频控制包位于数据切换点之后的第二行，每场发送一次。在紧随切换点之后的一行，没有辅助数据。切换点的位置由高清晰度信号的格式所决定，例如，在1125/60格式中，第8行上没有辅助数据。

音频控制包

音频控制包传送的是音频数据解码过程中要使用的附加信息，具有类似于标准清晰度的数据包结构，如图28所示。音频控制包包含有如下信息。辅助数据标志

（三个字）的数值为000_h、3FF_h、3FF_h。DID（一个字）的数值如表7所示，它用于区分相应的音频数据组。DBN的数值总为200_h而DC数值总为10B_h，它们均为一个字。UDW共有11个数据字，其结构可划分为五种不同的数据类型。音频帧号数据（AF）给出了视频帧的序列号，在每帧的音频样值数不为整数的情况下，它有助于识别音频样值的位置。RATE（一个字）用于指示音频数据的取样频率以及该数据处于同步还是异步。ACT字指示数据组内的有效通道数。DELM-n共6个字，用于指示音频处理对视频的相对时延的累积数，该视频为每通道对1&2和3&4音频样值间隔中的被测值。

这略不同于标准清晰度中的使用格式。RSRV（两个字）为保留字，以备后用。



► 图28. 音频控制包的数据结构

音频监视

► 应用注解

表 8. 音频数据的比特分配

Bit	UDW2	UDW3	UDW4	UDW5
B9	B8	B8	B8	B8
B8	偶校验	偶校验	偶校验	偶校验
B7	音频 ₁ 3	音频 ₁ 11	音频 ₁ 19	P ₁
B6	音频 ₁ 2	音频 ₁ 10	音频 ₁ 18	C ₁
B5	音频 ₁ 1	音频 ₁ 9	音频 ₁ 17	U ₁
B4	音频 ₁ 0	音频 ₁ 8	音频 ₁ 16	V ₁
B3	Z	音频 ₁ 7	音频 ₁ 15	音频 ₁ 23(MSB)
B2	0	音频 ₁ 6	音频 ₁ 14	音频 ₁ 22
B1	0	音频 ₁ 5	音频 ₁ 13	音频 ₁ 21
B0	0	音频 ₁ 4	音频 ₁ 12	音频 ₁ 20
Bit	UDW6	UDW7	UDW8	UDW9
B9	B8	B8	B8	B8
B8	偶校验	偶校验	偶校验	偶校验
B7	音频 ₂ 3	音频 ₂ 11	音频 ₂ 19	P ₂
B6	音频 ₂ 2	音频 ₂ 10	音频 ₂ 18	C ₂
B5	音频 ₂ 1	音频 ₂ 9	音频 ₂ 17	U ₂
B4	音频 ₂ 0	音频 ₂ 8	音频 ₂ 16	V ₂
B3	Z	音频 ₂ 7	音频 ₂ 15	音频 ₂ 23(MSB)
B2	0	音频 ₂ 6	音频 ₂ 14	音频 ₂ 22
B1	0	音频 ₂ 5	音频 ₂ 13	音频 ₂ 21
B0	0	音频 ₂ 4	音频 ₂ 12	音频 ₂ 20
Bit	UDW10	UDW11	UDW12	UDW13
B9	B8	B8	B8	B8
B8	偶校验	偶校验	偶校验	偶校验
B7	音频 ₃ 3	音频 ₃ 11	音频 ₃ 19	P ₃
B6	音频 ₃ 2	音频 ₃ 10	音频 ₃ 18	C ₃
B5	音频 ₃ 1	音频 ₃ 9	音频 ₃ 17	U ₃
B4	音频 ₃ 0	音频 ₃ 8	音频 ₃ 16	V ₃
B3	Z	音频 ₃ 7	音频 ₃ 15	音频 ₃ 23(MSB)
B2	0	音频 ₃ 6	音频 ₃ 14	音频 ₃ 22
B1	0	音频 ₃ 5	音频 ₃ 13	音频 ₃ 21
B0	0	音频 ₃ 4	音频 ₃ 12	音频 ₃ 20
Bit	UDW14	UDW15	UDW16	UDW17
B9	B8	B8	B8	B8
B8	偶校验	偶校验	偶校验	偶校验
B7	音频 ₄ 3	音频 ₄ 11	音频 ₄ 19	P ₄
B6	音频 ₄ 2	音频 ₄ 10	音频 ₄ 18	C ₄
B5	音频 ₄ 1	音频 ₄ 9	音频 ₄ 17	U ₄
B4	音频 ₄ 0	音频 ₄ 8	音频 ₄ 16	V ₄
B3	Z	音频 ₄ 7	音频 ₄ 15	音频 ₄ 23(MSB)
B2	0	音频 ₄ 6	音频 ₄ 14	音频 ₄ 22
B1	0	音频 ₄ 5	音频 ₄ 13	音频 ₄ 21
B0	0	音频 ₄ 4	音频 ₄ 12	音频 ₄ 20

系统化的 AES/EBU 音频

在节目制作和后期制作设施以及电视台中，串行数字视频和音频的应用已日益普遍。在许多情况下，视频和音频是作为共有的信号源；人们希望它们始终在一起并将它们作为一种数据流来处理。就应用实例而言，这样作的优点在于能在数字域中始终保持这种共有信号，并能通过一个串行数字视频路由切换器将它们一起切换。在一些特殊应用场合，人们又希望断开某些音频信号源，这时需要对数字音频进行解复用，而后通过一种AES/EBU数字音频路由切换器予以分别切换。在接收端，已复用的数字音频经过串行数字路由切换器之后，也许需要将音频信号

从视频信号中抽出以进行编辑、音频加工或作其它处理，这就要通过解复用器从串行数字视频信号取出AES/EBU音频信号。一般解复用器均具备串行数字视频的BNC端口和两路立体声对AES/EBU数字音频信号的连接器，可用以输出音频信号。

结语

由于视频图像得到愈来愈多的关注，而音频监视似乎渐渐被人所遗忘。然而，音频如同产品质量一样地重要。因此，必须继续坚持监视音频信号，使音频信号不致失真，甚至造成音频信号的丢失。

音频监视

► 应用注解

附录 A

通道状态协议

字节 0		
比特 0	0	通道状态块的一般应用
	1	通道状态块的专业应用
比特 1	0	正常音频模式
	1	非音频模式
比特 2-4	234	嵌入音频信号加重
	000	未指明的加重
	100	无加重
	110	50/15 微秒加重
	111	CCIT J.17 加重
		所有其它状态均保留以备后用
比特 5	0	未指明的默认锁定状态
	1	未锁定的源信号取样频率
比特 6-7	67	已编码的取样频率
	00	未指明的取样频率
	01	48KHz 取样频率
	10	44.1KHz 取样频率
	11	32KHz 取样频率
字节 1		
比特 0-3	0123	已编码的通道模式
	0000	未指明的模式
	0001	双通道模式
	0010	单通道模式
	0011	主 / 次模式
	0100	立体声
	0101	保留已定义应用
	0110	保留已定义应用
	0111	单通道双取样频率模式
	1000	单通道比取样频率模式 立体声模式 左
	1001	单通道比取样频率模式 立体声模式 右
	1111	多通道模式
		所有其它状态均保留以备后用

比特 4-7	4567	已编码的用户比特管理		
	0000	已指明的默认无用户信息		
	0001	192 比特块结构		
	0010	保留，用于 AES 18 标准		
	0011	已定义的用户		
	0100	符合 IEC 60958-3 的用户数据		
		所有其它状态均保留以备后用		
字节 2				
比特 0-2	012	附属取样比特的编码应用		
	000	最大音频取样字为 20 比特。 使用 AUX 未定义		
	001	最大音频取样字为 24 比特。 AUX 用于音频取样数据		
	010	最大音频取样字为 20 比特。 AUX 通道用于附加的音频数据通道(如对讲系统)		
	011	保留，用于用户已定义的应用 所有其它状态均保留以备后用		
比特 3-5	012	已编码的音频取样字长度		
	000	未指明的字长度		
		如果音频样值 数据为 24 比特	如果音频样值 数据为 20 比特	
	001	23 比特	19 比特	
	010	22 比特	18 比特	
	011	21 比特	17 比特	
	100	20 比特	16 比特	
	101	24 比特	20 比特	
			所有其它状态均保留以备后用	
	比特 6-7	01	校准电平指示	
00		未指明的校准电平		
01		按 SMPTE RP155 校准电平 (20dB 校准电平)		
10		按 EBU R68 校准电平 (18.06dB 校准电平)		
11		保留以备后用		

字节 3		
比特 7	0	未定义的多通道模式
	1	已定义的多通道模式
比特 0-6		通道数，当比特 7 为零时 通道数为比特值加 1
比特 4-6	456	多通道模式，当比特 7 为 1 时
	000	多通道模式 0
	100	多通道模式 4
	010	多通道模式 2
	110	多通道模式 3
	111	用户已定义的多通道模式
		所有其它状态均保留以备后用
比特 0-3		通道数，当比特 7 为 1 时 通道数为 1 加上该比特被认为是 二进制时的数值
字节 4		
比特 0-1	01	数字音频基准信号
	00	无基准信号
	01	1 级基准信号
	10	2 级基准信号
	11	在定义前保留且不使用
比特 2		保留
比特 3-6	3456	取样频率
	0000	未指明
	0001	24KHz
	0010	96KHz
	0011	192KHz
	0100	保留
	0101	保留
	0110	保留
	0111	保留
	1000	保留
	1001	22.05KHz
	1010	88.2KHz
	1011	176.4KHz
	1100	保留
	1101	保留
	1110	保留
	1111	已定义的用户

比特 7		取样频率比例标记
	0	无比例
	1	取样频率为已指明的 1/1.001 倍

字节 5	
比特 0-7	在作进一步定义前予以保留并设置为逻辑电平零

字节 6-9	
比特 0-7	字母数字通道原始数据

字节 10-13	
比特 0-7	字母数字通道目标数据

字节 14-17	
比特 0-7	本地样值地址码

字节 18-21	
比特 0-7	日期样值地址码的时间

字节 22	
比特 0-3	在作进一步定义前予以保留并设置为逻辑零电平
	如果随后的数据字节可靠，则随后标记设置为逻辑电平 1

比特 4	字节 0 至 5
比特 5	字节 6 至 13
比特 6	字节 14 至 17
比特 7	字节 18 至 21

字节 23	
比特 0-7	通道状态数据循环冗余校验

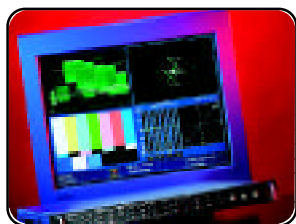
音频监视

► 应用注解



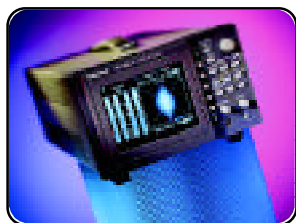
WFM700 系列 波形监视器

- 单个仪器即可监视和测量HD和SD信号
- HD和SD眼图的测量和抖动显示
- 可配置的模块式结构



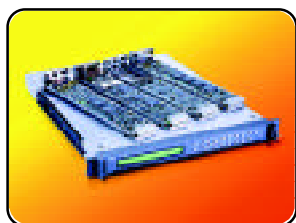
WVR600 系列波形监测仪

- 单个仪器即可监视SD、SDI和模拟分量信号
- 可选购的模拟和(或)AES/EBU音频监视
- 强大的FlexVu™显示功能，操作简单，使用灵活



764 音频监视器

- 两路平衡式或不平衡式的AES/EBU或SPDIF环通输入，四通道电平表显示
- 可调整的削峰/静音指示，支持多种技术参数



TG700 多格式视频发生器

- 多格式的模拟和数字测试信号发生器
- 可扩展的模块式平台

更多信息

泰克公司备有内容丰富的各种应用手册、技术介绍和其他资料，并不断予以充实，以帮助那些从事前沿技术研究的工程师们。

请访问 www.tektronix.com

泰克电子(中国)有限公司

北京市海淀区花园路4号
通恒大厦1楼101室
邮编：100088
电话：(86 10) 6235 1210/1230
传真：(86 10) 6235 1236

泰克上海办事处

上海市静安区延安中路841号
东方海外大厦18楼
邮编：200040
电话：(86 21) 6289 6908
传真：(86 21) 6289 7267

泰克广州办事处

广州市环市东路403号
广州国际电子大厦2107室
邮编：510095
电话：(86 20) 8732 2008
传真：(86 20) 8732 2108

泰克深圳办事处

深圳市罗湖区深南东路5002号
信兴广场地王商业大厦4302室
邮编：518008
电话：(86 755) 8246 0909
传真：(86 755) 8246 1539

泰克成都办事处

成都市人民南路一段86室
城市之心23层D-F座
邮编：610016
电话：(86 28) 8620 3028
传真：(86 28) 8620 3038

泰克西安办事处

西安市东大街
西安凯悦(阿房宫)饭店322室
邮编：710001
电话：(86 29) 769 1234-322
(86 29) 723 1794
传真：(86 29) 721 8549

泰克武汉办事处

武汉市武昌区民主路788号
白玫瑰大酒店924室
邮编：430071
电话：(86 27) 8789 3366-924
(86 27) 8781 2831
(86 27) 8731 8969
传真：(86 27) 8730 5230

泰克香港办事处

香港铜锣湾希慎道33号
利园3501室
电话：(852) 2585 6688
传真：(852) 2598 6260



© 美国泰克公司 (Tektronix, Inc.) 2003 年版权所有。 泰克公司的产品受正在申请或已批准的美国外国专利保护。 本手册之内容取代以前所有出版物的内容。 本公司保留随时更改技术规格和产品价格的权利。 TEKTRONIX 和 TEK 是泰克公司的注册商标， 本文提及之所有其它商业名称分别为其各自公司的服务标志、 商标或注册商标。 01/03 TD/XBS 21C-16463-1

Tektronix
Enabling Innovation