白皮书

麦克风拾音范围 – 影响音频性能的因素

9月 2026

目录

1 简介 3
2 物理录音范围 3
3 不同环境下的录制 4
4 不同环境下录制的测量结果 7
1 附录 8

1 简介

为音频录制设备明确规定一个确切的最大拾音范围往往并不简单，因为其性能与声学环境密切相关。虽然技术规格提供了一个基准，但实际录制范围取决于环境和声学因素之间的复杂相互作用，而这些因素因安装环境的不同而各异。

本白皮书探讨了决定麦克风实现可靠拾音距离的主要因素，并重点介绍语音清晰度作为主要参考标准。这些因素包括噪音等级、噪音的频谱特征、混响时间、麦克风的指向性，以及麦克风和录制设备的本底噪音。

为了建立一个切实可行的基准，我们以人声作为参考声源。噪音测量时会使用A加权滤波器，以使测量值与感知噪音水平之间具有更好的相关性。录制语音信噪比 (SNR) 的计算方法是：语音电平与A加权噪音电平之间的差值，单位为分贝（*有关SNR，请参见附录*）。实证测试表明，只要噪音在语音频率范围内具有相对平坦的频率响应，语音清晰度在信噪比 (SNR) 低至约0 dB时仍可被接受（即噪音和声音的强度相等）。

通过了解这些变量，安装人员和系统设计师在实际应用场景中选择和布置音频设备时，能够做出更明智的决策。

2 物理录音范围

麦克风的录音功能通常取决于环境背景噪音。限制物理录音范围的主要因素包括：

- **声压级**：并非所有声音传播的距离都相同。例如，枪声会产生极高的声压级，在距离一公里外仍能听见。不过，人声更轻，而且随着距离的增加，声音会更快地变弱。在估算实际拾音范围时，首先需要明确需要捕捉的声音类型，这样声源级才能确定可录制声音的上限。
- **背景噪音**：环境噪音越大，麦克风就越难将目标声音与其他声音分离并清晰捕捉。背景噪音因环境而异，并对目标声音直接产生干扰。其中包括风噪、交通噪音、机械噪音和人群噪音等。麦克风所在位置的噪音越高，其有效拾音范围就越短。这解释了为什么同一支麦克风在安静的走廊里和繁忙的十字路口，表现会大不相同。
- **噪音的频谱特征**：正如噪音的响度很重要一样，其所涵盖的频率同样重要。人类语音主要集中在中频段，大致在300 Hz到3,400 Hz之间。如果背景噪音集中在极低或极高频段，就较少会影响语音清晰度。然而，如果噪音与目标声音的频率范围重叠，即使在适中的噪音等级下，也很难将语音与噪音区分开来。
- **麦克风位置处的混响时间**：混响是指声源停止发声后，声音仍持续存在，这是由于声音从墙壁、天花板、地板及其他表面反射所致。较长的混响时间会降低语音清晰度，即使在整体声级和信噪比看似可接受的情况下也是如此。在回声非常强烈的空间（例如宽敞的空旷大厅或铺有瓷砖的房间）中，语音信号传到麦克风时会伴随无数次房间反射，导致声音变得模糊不清。这使听众和录音系统都难以清晰地分辨出单个词语。
- **麦克风设置的指向性**：不同麦克风对来自不同方向的声音的灵敏度各不相同。全向麦克风能从各个方向均匀地拾取声音，这也意味着它会从各个方向拾取背景噪音和混响。定向设置（例如用于波束成形的多个麦克风配置）可将灵敏度集中于特定方向，从而有效降低来自其他角度的噪音和混响的影响。这可以在嘈杂或混响环境下显著延长麦克风阵列的有效拾音范围，但仅限于麦克风阵列指向的方向。
- **麦克风和录制设备的本底噪音**：每支麦克风和每台音频录制设备都会产生少量内部电噪音，这被称为本底噪音。这设定了系统可捕捉声音的下限。无论环境多么安静，低于本底噪音阈值的声音都无法被清晰地录制下来。在比较麦克风时，本底噪音规格越低，通常意味着该麦克风能更好地捕捉更微弱的声音；如果噪音等级和混响时间足够低，这可能会带来更远的拾音范围。

为了简化分析，我们可以从一个理想化的场景入手，该场景中没有背景噪音、没有反射声，也没有混响，并将人声作为参考声源。在这种情况下，关键问题在于：我们应该将多少信噪比 (SNR) 作为可接受拾音范围的阈值?

实证测试给了我们明确的答案。当我们将A加权滤波器应用于系统噪音，并将噪音的平均均方根 (RMS) 电平与语音的平均RMS电平进行比较时，只有当信噪比 (SNR) 降至0 dB以下时，语音清晰度才会开始下降。因此，我们将0 dB信噪比 (SNR) 作为录制人声的限制标准。只要噪音的频率响应相

对平坦（类似于白噪音），且处于语音频率范围内，这一结论就成立；而麦克风本底噪音通常就符合这一条件。

例如，当1 kHz的音调以94 dBspl播放时时，数字麦克风的灵敏度可能为–36 dBFSrms（*有关dBFSrms，请参见附录*）；根据麦克风数据表，此处将数字化满量程正弦波定义为0 dBFSrms（*有关dBspl，请参见附录*）。当麦克风增益设置为0 dB时，这适用于许多内置数字MEMS麦克风的安讯士产品。当您将此麦克风放置在背景噪音等级低于10 dBspl(A) 的非常安静的环境中时（*关于dBspl (A)，请参见附录*），麦克风内部的噪音就会成为主要噪音源。在此例中，麦克风数据表规定其数字输出为–105 dBFSrms(A)。该麦克风的信噪比 (SNR) 是参考电平与底噪之间的差值：−36 − (−105) = 69 dB。

我们将这一电噪音底限转换为等效声学噪音等级：94 − 69 = 25 dBspl(A)。

现在，我们来计算在0 dB SNR条件下，能够录制人声的最大距离。根据国际电信联盟 (ITU) 标准，人声在1米处的标称平均语音电平 (ASL) 约为57 dBspl。在自由空间中，无反射物体时，声压级随距离减小，遵循平方反比定律：

$$dB = 20\log\frac{reference\ distance}{test\ distance}$$

。

如果参考距离为1米，要使该麦克风达到0 dB SNR，语音电平必须降至25 dBspl(A)；这比1米参考电平降低了32 dB。重新排列公式得到：

$$test\ distance = \frac{reference\ distance}{10^{\frac{-(57-noise\ level)}{20}}} = \frac{1}{10^{\frac{-32}{20}}} = 39.8\text{m}$$

。

这意味着在自由空间中，在没有背景噪音、没有风声且没有混响的情况下，该麦克风理论上可以在约40米处录制清晰的语音。这是其理论上的最大拾音距离。

然而，环境因素对噪音等级的影响比内置麦克风的噪音底限更大。下班后，通风系统仍在运转的安静办公室，其噪音等级约为28 – 32 dBspl。仅此一项就可能使有效拾音范围减少约50%，最大拾音距离降至约20米。如果我们在城市环境中进行户外录音，将会出现多种噪音的混合：风噪、交通噪音、人类活动产生的噪音等。在使用全向麦克风时，每个噪音等级和噪音频谱都会有其对应的拾音范围。在此处，如果交通噪音较大，最大拾音范围可达2 – 4米。

一个实际需要考虑的问题是，在0 dBSNR阈值或接近该阈值时录制的人声，需要显著的数字增益才能被听到，因为录制的人声电平非常低，否则在播放时会低于我们自身的听觉阈值。在下面的录音示例中，音频文件已被放大器放大+40 dB，以补偿远处人声电平。

3 不同环境下的录制

本节详细介绍了录制地点，并附有视频录制的截图以及在不同距离下采集的语音样本。

示例1：地点位于户外，有来自高速公路的交通噪音。背景噪音等级在62至68 dBspl[A]之间，还伴有一定的风噪。



图 3.1 *录制地点靠近一条高速公路（见红色圆点）。*



图 3.2 *在录制地点，安讯士穿戴式摄像机的最远可录制距离为4米。*

建议在收听录音时佩戴一副耳机。

示例2：一条安静的走廊，背景噪音等级为32 dBspl[A]。在这里，走廊可以被看作是一种长声波波导。声音在走廊中传播，在反射面上多次反射后，到达麦克风时的声压级比在自由空间中相同距离时要高。

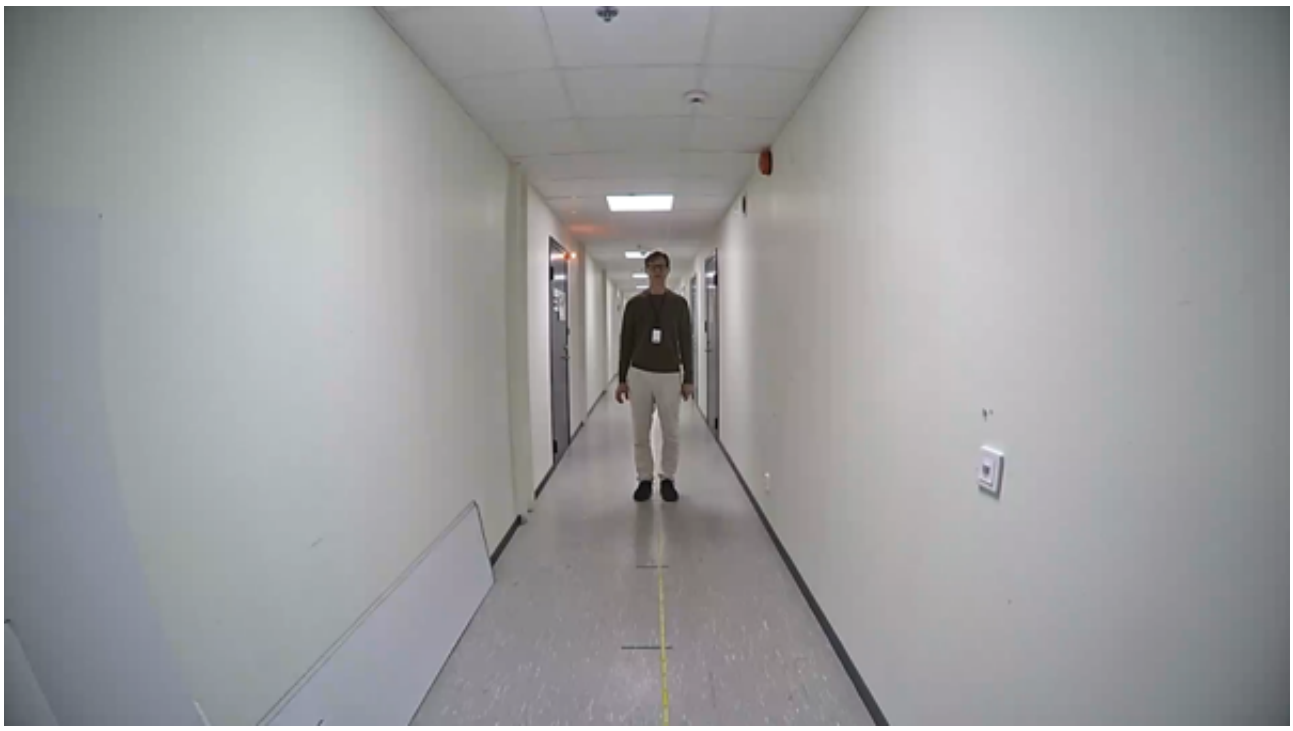


图 3.3 *最远可录制距离为27米。*

建议在收听录音时佩戴一副耳机。

示例3：一个宽敞的大堂，背景噪音等级为42 dBspl[A]。在这里，墙壁上的扩散板后面安装了吸音材料，从而消除了大部分早期反射声。从长距离来看，声音沿着地板与反射天花板之间的声学通道传播。



图 3.4 *在最远可录制距离40米处，扬声器几乎看不见，但声音依然能听见。*

建议在收听录音时佩戴一副耳机。

4 不同环境下录制的测量结果

本分析主要侧重于语音拾取。像枪声这样更响亮的声音，在更远的距离也能听见。我们对音频录制中的每句话都进行了分析和测量，并估算了每句录音的信噪比 (SNR)。

信噪比 (SNR) 的估算公式为：语音+噪音-背景噪音。更准确的方法应该是计算语音电平-噪音电平，但无法在不考虑噪音的情况下单独测量语音电平。这意味着当语音电平接近噪音电平时，信噪比 (SNR) 的准确度会降低。

地点和背景噪音：	有交通噪音的户外，61–68 dBspl[A]	在安静的走廊里，32 dBspl[A]	在宽敞的大堂内，42 dBspl[A]	自由空间，无反射，完全静默
距离（米）	**信噪比 [dB]（估计值）**	**信噪比 [dB]（估计值）**	**信噪比（[dB]，估计值）**	**理论信噪比（SNR）= 语音级麦克风噪音 [dB]**
1.	2.4	29.4	21.5	32.0
2	1.0	25.6	17.2	26.0
4	0.5	23.4	14.2	20.0
8	0.4	18.9	11.2	13.9
16	听不见	17.5	6.3	7.9
24	听不见	14.8	5.4	4.4
32	听不见	走廊太短	4.6	1.9
40	听不见	走廊太短	4.7	0.0

理论上，当距离加倍时，语音电平应降低–6 dB。这也意味着信噪比 (SNR) 应相应降低。从这张表中可以看出，走廊和大堂的信噪比 (SNR) 都比预期的要好。原因是房间内的反射导致声压级升高。特别是在走廊里，墙壁、地板和天花板都会充当声音的波导。这使得在走廊里、距离24米处进行测试时，其响度和信噪比比自由空间中的理论信噪比高出约10 dB。

每当背景噪音超过单个全向麦克风的声学等效本底噪音时，拾音范围就完全取决于环境背景噪音，而非麦克风本身。

附录 1　　附录

- **分贝声压级 (dBspl)**：声压级以20 μPa为基准，单位为分贝。在测量噪音和语音电平时，通常使用时间积分均方根电平。这有时也被称为“电平”。

$$RMS = \sqrt{\frac{1}{T2-T1}\int_{T1}^{T2}[f(t)]^2\,dt}$$

最常见的两种积分时间分别是时间窗口为125毫秒的“快速RMS”和时间窗口为1秒的“慢速RMS”。在噪音底限分析中，通常会对该电平取较长时间内的平均值。

- **经A加权滤波器加权的声压级 (dBspl[A])**：有时也写作dB(A)。这是dBspl电平的A加权值。A加权是一种对麦克风信号进行滤波的方法，旨在使测量得到的声压级与人耳听感更吻合。这是因为人耳对低于500 Hz的低频和高于10 kHz的高频的敏感度较低。

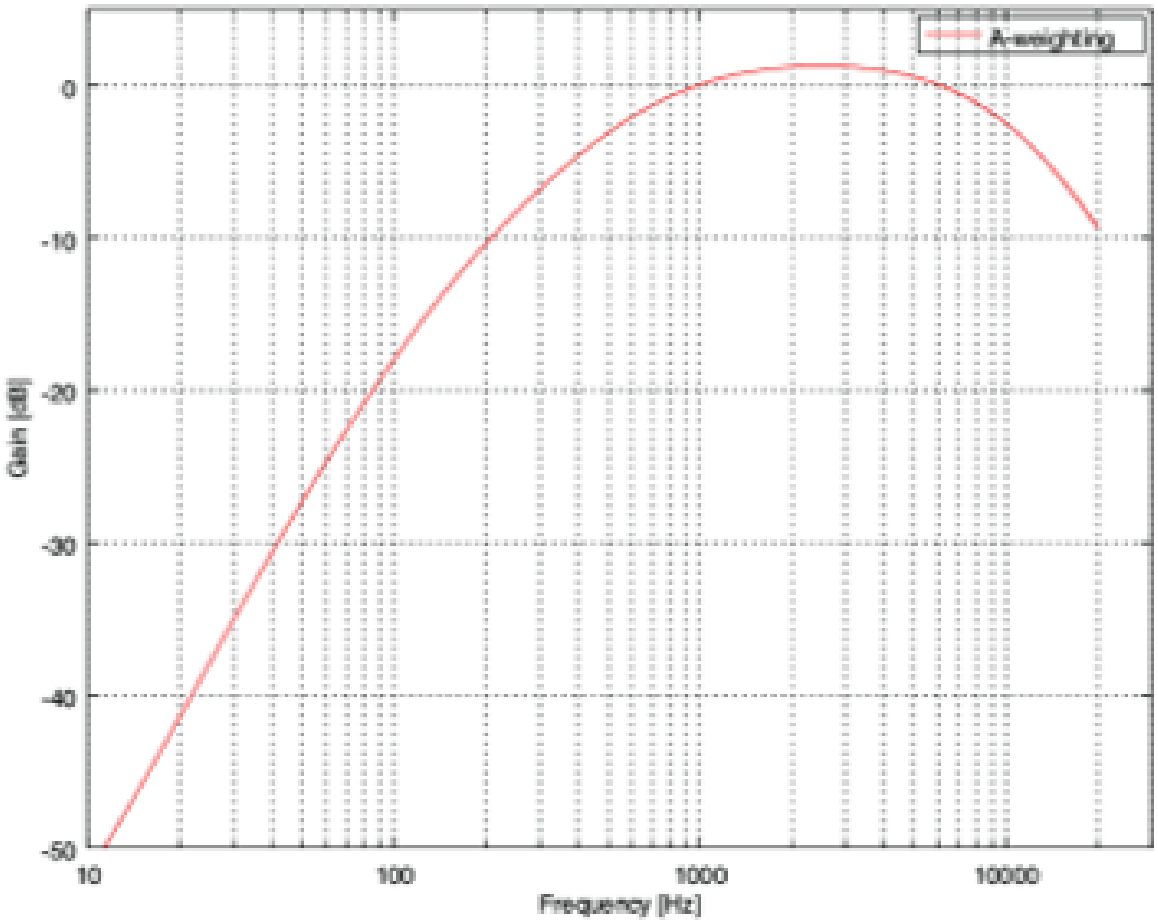


图 .1 *A加权滤波器*。

- **满量程分贝峰值 (dBFSpeak/dBFSp)** 是指以可用的最大数字表示值（满量程）为基准的数字音频振幅。无削波的最大数字振幅正弦波，峰值为0 dBFSp。

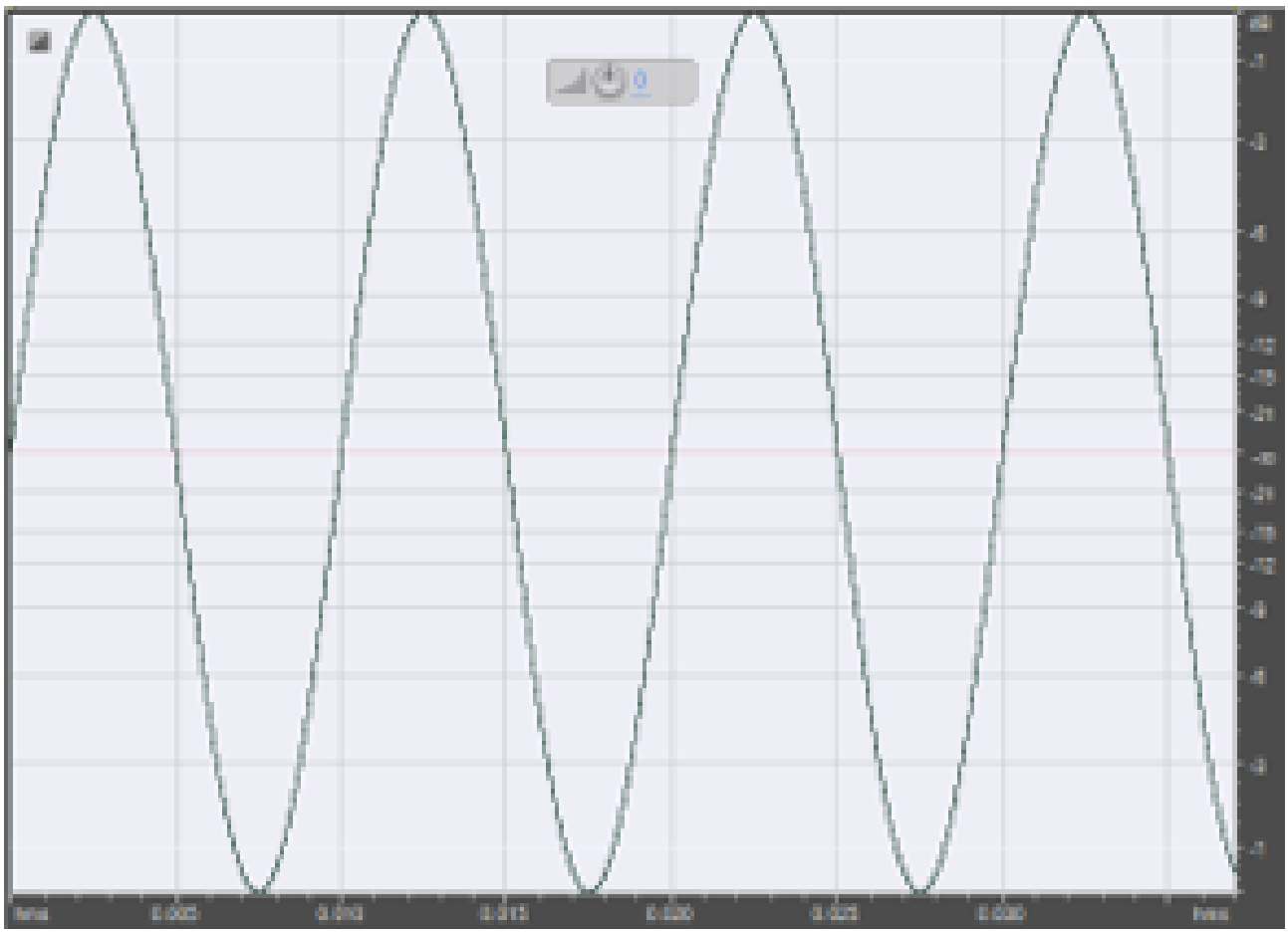


图 .2 *一个振幅为0 dBFSpeak的信号示例。*

- **dBFSrms**指数字信号的数字音频电平。一个0 dBFSpeak的全尺度正弦波，其均方根电平会因单元的定义方式不同而有所差异。在麦克风数据表中，满量程正弦波被定义为 0 dBFSrms，这也是本白皮书采用该定义的原因。总体而言，全尺度正弦波的另一种定义–3 dBFSrms，在数学上更为准确，也更符合模拟域中均方根电平和峰值电平的定义方式。根据这一另一种定义，一个满量程方波为 0 dBFSrms。
- **麦克风灵敏度**是指当声学输入为1khz正弦波（1pa或0db a或94dbspl）时麦克风的输出电平（同一声压级有三种表达方式。）
- **信噪比 (SNR)**：当音频电平以dB表示时，该比率计算为信号电平与噪声电平之间的差值。
- **麦克风SNR**是通过计算94 dBspl的1 kHz测试音与麦克风在完全静默环境下的本底噪音之间的输出电平差值来得出的。数字MEMS麦克风的典型信噪比 (SNR) 为69 dBspl[A]。
- 本白皮书中所指的**语音信噪比 (SNR)** 是指人类语音与无语音时的A加权噪音电平之间的比较。信噪比 (SNR) 的单位是分贝。在计算语音的理论拾音范围时，我们使用0 dB语音信噪比 (SNR) 作为最低可接受信噪比的定义。在此背景下，噪音可能来自环境噪音、电气噪音以及麦克风本底噪音。
- **麦克风噪音电平**麦克风的数字噪音电平以dBFSrms为单位进行测量。通常，最好的数字麦克风的理论噪音底限为–105 dBFSrms[A]。
- **麦克风的声学等效噪音**可按以下公式计算：94 dBspl（测试电平）– 麦克风信噪比。一个常见的例子就是94 – 69 = 25 dBspl[A]。

T10245238_ZH/ZH/M1.4/2609

关于安讯士 (Axis Communications)

安讯士通过打造各种解决方案，提高安全水平和企业效益，旨在创造一个高度智能、更加安全的世界。作为一家网络技术公司和行业领导者，安讯士致力于推出视频监控、访问控制、内部通信和音频系统解决方案。安讯士通过智能分析应用程序增强解决方案，并提供高质量培训支持。

安讯士在全球50多个国家和地区设有办事机构，拥有超过5,000名尽职的员工，并与遍布世界各地的技术和系统集成合作伙伴携手并进，为客户带来高价值的解决方案。安讯士创立于1984年，总部位于瑞典。

© 2026 — Axis Communications AB. AXIS COMMUNICATIONS、AXIS、ARTPEC和VAPIX是Axis AB在不同司法辖区内的注册商标。所有其它商标均为其各自所有者的财产。我们保留在未预先通知的情况下提出修改的权利。