

高分辨率声音处理策略

作者：Jill B. Firszt, 博士

美国威斯康辛医学院

HiResolution (HiRes) 高分辨率声音处理策略可由 HiResolution 高分辨率仿生耳系统实现，并通过 SoundWave 调机软件进行编程调试。能够采用高分辨率声音策略的声音处理器包括 CII 耳背式声音处理器, HiRes Auria 耳背式(BTE) 声音处理器以及白金系列体佩式声音处理器 (PSP)。高分辨率声音处理策略利用最新最先进的声音处理技术，已经成为 HiResolution 仿生耳系统的必要组成部分。

人工耳蜗的声音处理

声音处理方式决定了某种人工耳蜗系统将声音转化为针对听觉神经的电刺激时所依据的方法。所有声音处理策略都由下面几个方面组成：1) 从麦克风采集的声音输入信号怎样被分析并转换为电刺激信号；2) 电流在何时以何种形式被传递至电极刺激点；3) 在每个电极刺激点刺激波形的形式。声音处理程序的最终目标是传递用来精确表达声音的电刺激，使人工耳蜗使用者感知声学信号的方式接近正常听力的人，即便患有重度或极重度耳聋。

声学信号的精确的电刺激表达对于理解言语至关重要，特别

是在困难听音环境中。语音是一种复杂的声学信号，包含频率，强度，和时间信息。人工耳蜗系统对这些信息进行编码的精确度各不相同。语音信息中有很高的冗余度，所以在安静环境中，如果缺失某些信息，语音仍然可以被理解。但是，在困难环境中，声音表达的精确度变得更重要，因为非常多的信息会被噪声掩盖。

例如，如图 1 所示音节 /ba/ 和 /bi/ 频域与时域波形。两种信号大体的时间波形包络非常相似，但是它们的频谱却不同。与这种情况相反，图 2 所示音节 /ba/ 和 /wa/ 的频域与时域波形。这对信号的频谱相似，但是时间波形的包络却不同。为了使得人工耳蜗使用者能够分辨图 1 中的信号，人工耳蜗系统必须能够精确表达频域信息。另一方面，为了能够分辨图 2 中的信号，人工耳蜗系统就必须能够精确表达时域信息。因此，人工耳蜗系统必须在时域和频域都能以高精度度表达信号。

高分辨率声音处理策略的设计目标是采用高精度的电刺激来表达传入的声学信号，并传递给植入耳蜗的电极，而且更精确地表达时域和频域信息。

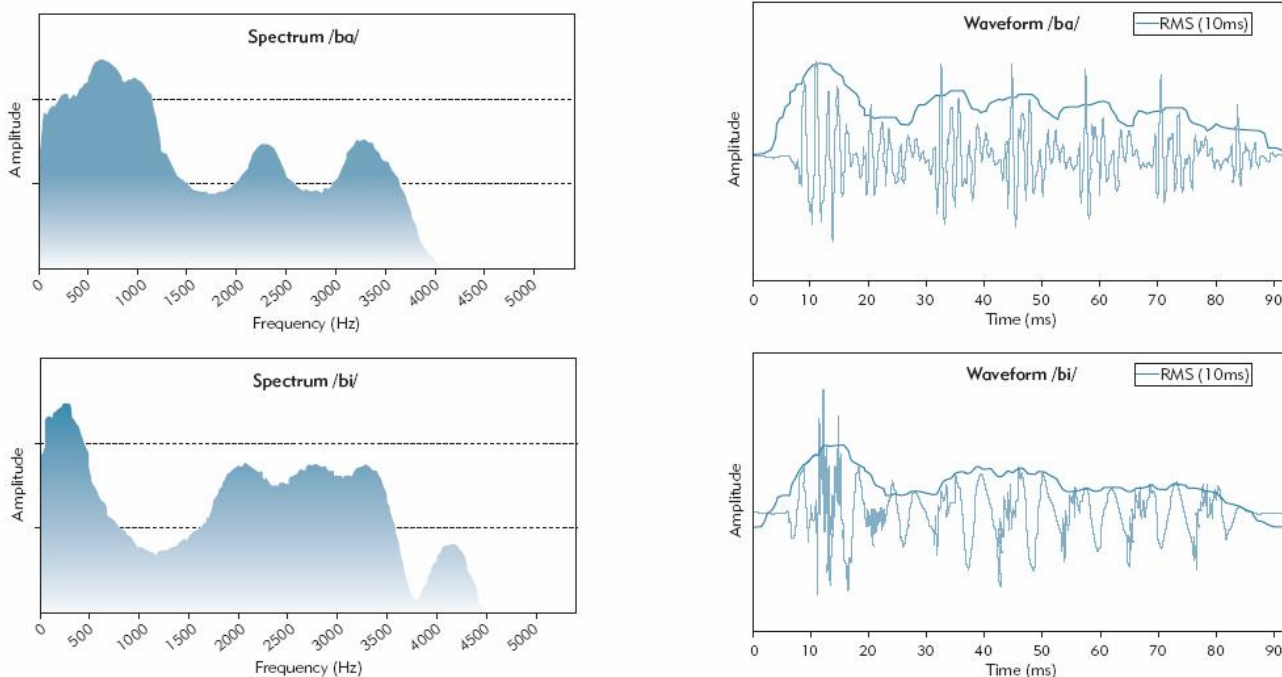


图 1. 音节 /ba/ 和 /bi/ 的频谱和时间波形。

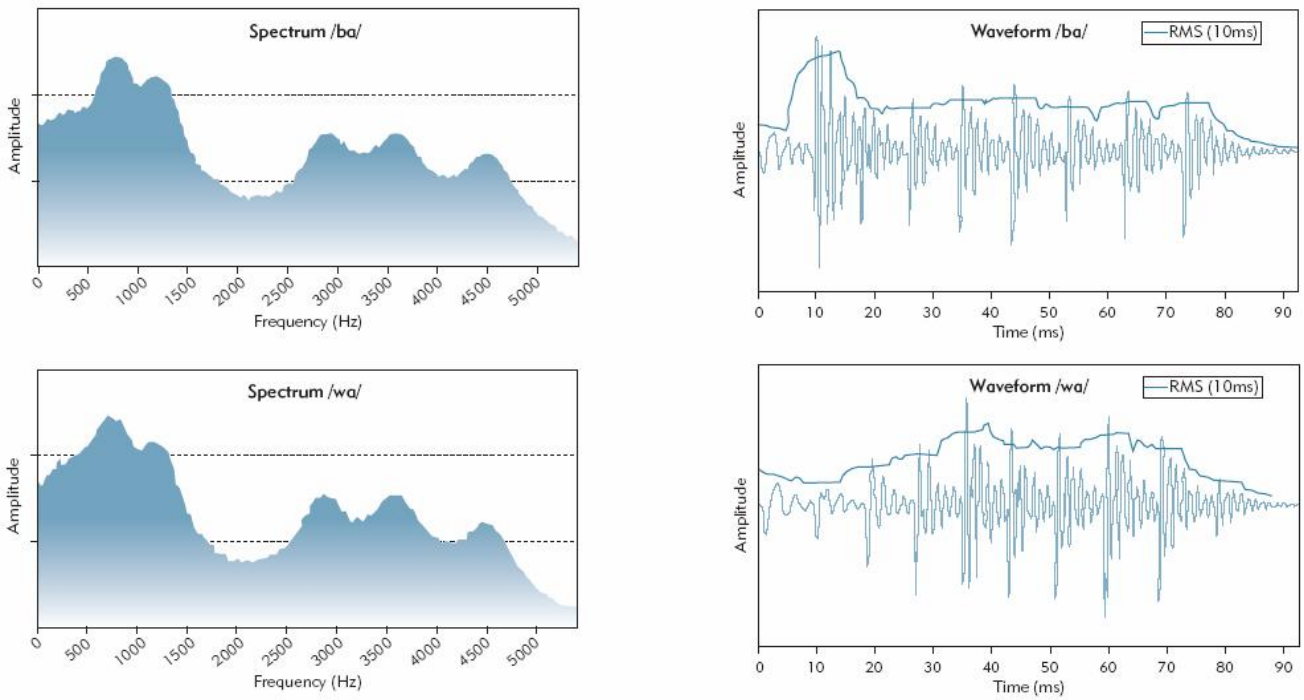


图 2. 音节 /ba/ 和 /wa/ 的频谱和时间波形。

高分辨率声音处理模式

为了更好地帮助理解高分辨率声音处理策略，我们将介绍一下声音被人工耳蜗系统获取和将电刺激传递给植入耳蜗的电极的过程。这些步骤包括：

1. 声音窗口
2. 频率映射
3. 波形分析
4. 刺激波形
5. 刺激速率
6. 电流输出源

声音窗口： 声音窗口即允许进入人工耳蜗系统的声音输入强度的范围。具备正常听力的人的听力动态范围可达 120 分贝（以声学听觉阈值为基准）。相对照而言，人工耳蜗使用者的电刺激信号的动态范围非常狭窄，在 10 到 20 分贝之间（以电刺激听觉阈值为基准）。因此，人工耳蜗系统必须将声学信号进行变换，才能保留重要信息，并将其编码以适应电刺激的狭窄动态范围。

高分辨率策略解决人工耳蜗使用者狭窄动态范围的方法有两种：1）扩展进入人工耳蜗系统的声音动态范围（输入动态范围，IDR）；2）采用双动作自动增益控制（AGC II）。系统的麦克风瞬时输入动态范围是 84 分贝，完全能够包含变化范围在 50 分贝的言语声音信号的波峰和波谷。在系统中，动态输入范围（IDR）的可编程范围是 20-80 分贝。如果 IDR 太狭窄，

低于 IDR 范围的声音会被丢弃，高于 IDR 的声音根据处理器的参数设置被斩断或被压缩。AGC II 包含一个增益由两个电压控制决定的自动增益放大器。慢动作控制的压缩阈值为 57dB SPL，启动时间为 325ms，恢复时间为 1000ms。第二个控制为快动作方式，压缩阈值较高为 65dB SPL，启动时间小于 0.6ms，恢复时间为 8ms。通常慢动作控制决定日常听音环境下的放大器增益，快动作控制则在对短暂的响亮声音作出反应时启动。

这样系统就能够在使用者电刺激范围内将范围宽广的言语信号强度进行编码，而不需要处理器的调节。较宽的输入动态范围使高分辨率策略的使用者听到不同距离，不同说话人说出的强度变化范围很宽的言语。图 3 显示了改变输入动态范围时对词语和句子识别能力的影响（Allen and Donaldson 2000）。对于 70dB 的输入强度，IDR 没有影响。对于 60dB 的输入强度，辅音识别能力开始受到影响。在 50dB 时，IDR 对辅音和句子的言语识别能力的影响更大。简而言之，总体言语强度减小时，对于辅音和句子的识别来说，更需要较宽的动态输入范围。真实世界中的声音强度，会因为说话人的音量，或同人工耳蜗使用者的距离等这些因素，产生非常大的变化，所以较宽的输入动态范围对于听懂言语非常重要。

频率映射： 声学信号进入人工耳蜗系统后，便被转化为数字信号，通过一组带通滤波器，通常范围在 250 到 8000Hz（正常人能听到 20 到 16000Hz 的声音）。在高分辨率仿生耳系统中，共采用 16 个按照对数比例划分的带通滤波器。每个滤波

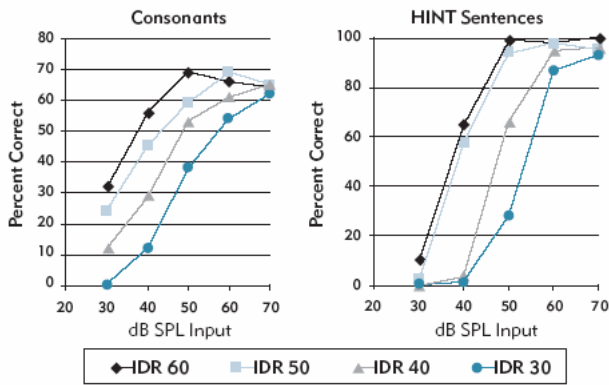


图 3. 四种 IDR 情况下辅音和句子的识别能力与声音强度的关系。

器的输出信号（经过进一步处理后）传送到植入耳蜗的电极刺激点。较低频率的滤波器输出信息送至靠近耳蜗顶部的刺激点，而较高频率的滤波器输出信息送至靠近耳蜗底部的刺激点。这样，频率信息被传递到耳蜗的不同部位，仿照正常人耳的音位映射的关系。

波形分析：频率映射提供了声音所包含的重要的频率信息（即各频率信号的幅度），然而每个频率段中包含的时间信息对于言语理解和音乐的欣赏（Shannon et al. 1995, Van Tassel et al. 1992）来说也非常重要。根据 Rosen（1989 年）的研究结果，言语的时域结构主要由三个参数构成。他认为，“包络”信号为 2-50Hz，提供韵律信息（重音，音节划分）。“周期”信号在 50-500Hz，表示音节信息，比如辅音种类和发音方式，以及声调信息。时域的“精细结构”信号范围在 600 – 10000Hz，包含有关辅音发音位置和元音特性的信息。

美国 Advanced Bionics 公司以前的人工耳蜗系统能够在 8 个输出滤波器上对 200-400Hz 的时间变化信息进行编码。但是，HiRes 采用不同的方法，能够在 16 个滤波器上表达高达 2800Hz 的时域变化。因此系统还原原始信号中时域信息的速度比以往快几个数量级，从而能够提供“精细结构”信息，提高声音理解能力，特别是在噪声环境中和音乐理解方面。

信号的时间变化信息从 16 个滤波器中提取后，输出的波形用来调制一个脉冲序列。调制后的脉冲序列传递到耳蜗内排列的相应电极刺激点。

刺激波形和速率，电流输出源：高分辨率仿生耳系统能够在电极刺激点上输出脉冲和模拟波形，但是当前只采用脉冲波形和单极耦合方式。脉冲宽度范围在 11 到 229 微秒/相。每个刺激点上每秒输出的脉冲数目（pps）同脉冲宽度，通道数目（1-16 个），以及脉冲以单个次序发送还是在两个刺激点同时

（成对）发送都有关系。系统具备 16 个独立的电流输出源，使其能够实现两个或多个电极刺激点同时发送电流。高分辨率声音策略能够根据使用的电极数目，调整参数使刺激速率最大，并根据病人需要调整脉冲宽度。对于 16 个电极采用单个脉冲刺激方式和窄脉冲时，刺激速率为每个通道 2900 脉冲/秒。对于 16 个电极采用成对刺激方式和窄脉冲时，刺激速率可超过每个通道 5100 脉冲/秒。

只有将精细的时域波形和快速的刺激速率相结合才能为人工耳蜗使用者提供高分辨率的信号。图 4 显示了采用低分辨率时域波形和缓慢刺激速率时，最终电极输出的刺激波形。与之对照，图 5 显示采用高分辨率时域波形和快速刺激速率时，最终电极输出的刺激波形。图 5 中的刺激波形才能最精确地表达原始的信号波形，并在电极刺激点上输出。

高分辨率声音处理技术，专业人员，以及人工耳蜗候选者

对于人工耳蜗专业人员来说，了解技术构成和病人使用体验后，会对高分辨率声音处理模式有更深入的理解。16 个电极刺激点，16 个独立电流输出源，双动作自动增益控制，和宽广的动态范围等都需要了解。另外，完整频谱的表达与时域包络的精确提取，高速的刺激速率等都是高分辨率策略的特点。可以这么说，高分辨率策略结合了同步模拟刺激策略（SAS）和连续间隔采样（CIS）的优势，即完整频谱信号分析，更精确的时域波形表达和高速脉冲刺激。

在向人工耳蜗候选者或者植入者介绍这些信息时，需要简要解释声音信号处理的方法。对于有些病人来说，利用“更新的”，“更快的”和“更好的”这些概念，配合说明高分辨率策略可以在耳背式处理器上使用，就能满足要求。对于其他病人来说，可能需要解释讨论包络跟随和刺激速率这些概念。专业人员和病人了解高分辨率对于日常生活带来的帮助后会加深理解。临床研究结果表明，曾经使用过传统声音处理策略和高分辨率策略的病人，经过比较后，采用高分辨率策略后的言语识别能力显著提高，并且大多数病人更喜欢高分辨率策略（Koch et al. 2004）。

总之，高分辨率仿生耳系统的先进信号处理能力能够为人工耳蜗植入者提供声音信号高精确度的表达方式。言语识别效果同传统的声音处理策略比较起来有了很大改善，并且大多数病人更喜欢听高分辨率策略提供的声音。我们不能忽视病人的自身因素（例如，对于成人来说，耳聋时间长短，残余听力；对于儿童来说，植入的年龄，交流方式等）对人工耳蜗的效果影响。然而，我们也不能忽视技术革新对于病人效果的贡献。而不论病人的术前的听力和交流情况如何，提高

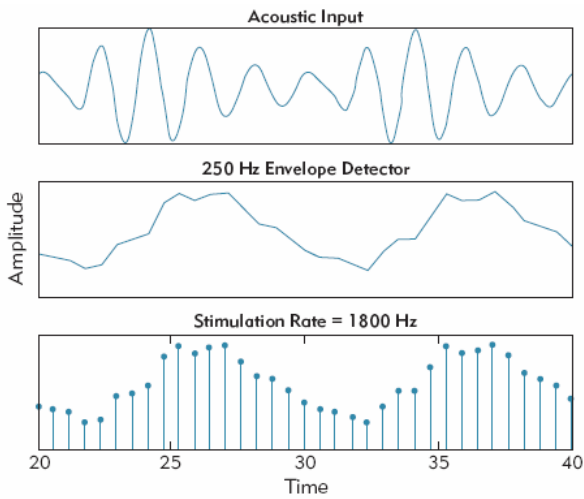


图 4. 声音刺激信号，时域波形提取，和刺激速率：250Hz 时间分辨率和 1800 脉冲/秒刺激速率（低分辨率，低刺激速率）。

原始声音信号的电刺激表达方式，将会提高病人改善听觉的机会。高分辨率策略向改进信号表达方式，并随之最终改善病人效果迈出了关键的一步。关于高分辨率仿生耳系统更多声音处理能力的进一步研究仍在开展过程中。

参考文献:

- Allen SL, Donaldson GS. (2000) unpublished data.
 Koch DB, Osberger MJ, Segel P, Kessler D. (2004) HiResolution and conventional sound processing in the HiResolution Bionic Ear: using appropriate outcome measures to assess speech recognition ability. *Audiol Neuro-Otol* (in press).

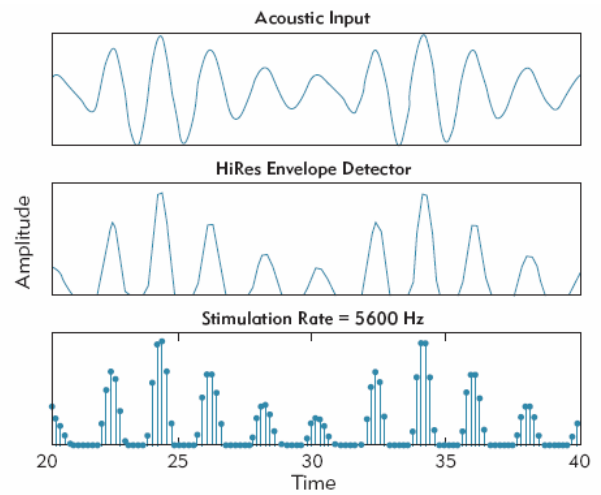


图 5. 声音刺激信号，时域波形提取，和刺激速率：HiRes 时间分辨率和 5600 脉冲/秒刺激速率（高分辨率，高刺激速率）。

Rosen S. (1989) Temporal information in speech and its relevance for cochlear implants. In *Cochlear Implant: Acquisitions and Controversies*, ed. B. Fraysse, N. Couchar, pp. 326.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. (1995) Speech recognition with primarily temporal cues. *Science* 270:303304.

Van Tasell DF, Greenfield DG, Logemann JJ, Nelson DA. (1992) Temporal cues for consonant recognition: training, talker generalization, and use in evaluation of cochlear implants. *J Acoust Soc Am* 92(3):12481257.