

人工智能应用

长期以来,公司大力支持产品技术研发和创新,且不断鼓励同事持续地技术创新和优化公司的产品,提高研发水平,增强企业的竞争力,公司对每一项改良、优化、新技术,创造发明都会进行奖励和表彰,项目主要是产品上的创新应用,无线电和声学上的理论设计应用。

TWS (True Wireless Stereo) 真无线立体声耳机,其外部完全摒弃了线材连接的方式,与传统蓝牙耳机相比可以实现单只耳机连接手机使用、两只耳机组成立体声播放的效果,不仅具备听歌看视频、打游戏、语音通话的能力,部分TWS耳机还带有运动状态监测、体温检测、心率检测等功能,因其体积小巧、功能丰富等特点,TWS耳机蕴藏着多媒体入口的巨大潜力,这几年受到市场的热烈追捧。TWS耳机得以诞生和快速发展的核心推动力是蓝牙连接技术的发展,更高的带宽、更快的传输速度、更远的传输距离以及更低的功耗有力地推动了行业兴起。作为TWS耳机的核心元器件,蓝牙音频SoC承担着无线连接、音频处理、电源管理、智能交互等功能,对于产品体验起重要影响。当中带出支援清晰音质,支持环境降噪(ENC)及Hybrid主动降噪功能,并进一步搭载最新降噪算法,为蓝牙耳机提供丰富及更强大的效果。PESQ是一种客观的评估方法,用于测量用户体验到的语音质量。以下是一些传统算法的效果。

	Noisy	LogMMSE [2]	EM_LogMMSE [3]	SEGAN [1]
SDR (1dbSNR)	1.189	6.541	7.493	10.244
SDR (3dbSNR)	3.16	7.385	8.51	11.349
SDR (6dbSNR)	6.085	8.615	9.979	12.689
PESQ (1dBSNR)	1.744	2.161	2.247	2.319
PESQ (3dBSNR)	1.873	2.283	2.373	2.454
PESQ (6dBSNR)	2.068	2.46	2.51	2.61

表 1. 不同降噪算法,对比的实验结果(越高越好)

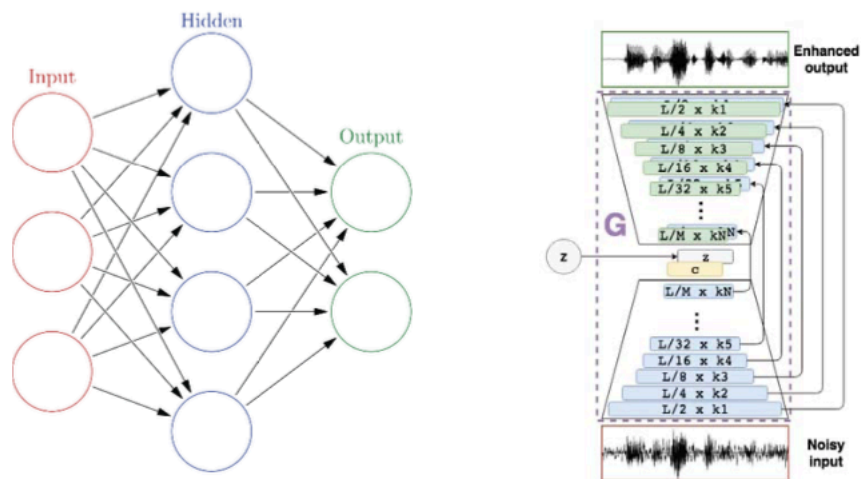
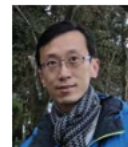


图 1. (左) 一个小神经网络其中包括：输入层、隐藏层和输出层。(右) .SEGAN[1] 的人工神经网络设计

在AI方面，目前较为先进的技术是利用人工智能网络对带有噪音的语音进行降噪。人工神经网络由多组称为人工神经元的连接单元或节点组成，其中包括输入层、隐藏层和输出层。在大型神经网络中，通常由多个隐藏层和一个输入及输出层组成。图 1(右)是SEGAN的多层人工神经网络的设计，通过输入带有噪音的语音。输入层传达语音到隐藏层，隐藏层会提取语音的特征(蓝色和绿色的方格)，例如音韵，谐波等等。这些特征会在隐藏层向上传递，会用作重新建构一段较清晰及噪音较少的语音。

撰文：Eric（工程部副总）



随着IC芯片设计和晶圆代工技术的不断进步，促进了上述技术的发展。将来在耳机上，AI应用会愈来愈普及。

1. S. Pascual, A. Bonafonte and J. Serra, "SEGAN: Speech enhancement generative adversarial network," arXiv preprint arXiv:1703.09452, 2017.
2. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," IEEE transactions on acoustics, speech, and signal processing, vol. 33, p. 443-445, 1985.
3. D. P. K. Lun, T.-W. Shen and K. C. Ho, "A novel expectation-maximization framework for speech enhancement in non-stationary noise environments," IEEE/ACM transactions on audio, speech, and language processing, vol. 22, p. 335-346, 2013.