

184th Meeting of the Acoustical Society of America

8 – 12 May 2023

Place:

Chicago, Illinois/USA

Title:

Implementing neural networks in low-latency audio applications

Authors:

Thiago Lobato, Roland Sottek, Michael Vorländer* (*IHTA)

Abstract:

The use of neural networks is becoming increasingly prevalent due to their ability to represent complex relationships and solve complex problems. However, implementing these models in systems that require low-latency output can be challenging, especially for practitioners who are used to developing their models in controlled environments like Python notebooks. Another issue is the high computational cost of complex models, which limits the minimum possible latency. This paper presents approaches for deploying models in audio applications, discusses the advantages and disadvantages of each approach, and presents strategies to reduce the inference cost of models without significantly sacrificing accuracy, using techniques such as model quantization. To illustrate these methods, example implementations of real-time beamforming deconvolution and real-time music DSP processing are shown.

Find more event abstracts in our [>> abstracts archive <<](#)