

## Application Brief

# 使用边缘 AI 和 Sitara 处理器进行关键字检测



Reese Grimsley and Tarkesh Pande

Sitara MPU

时间序列信号是嵌入式系统中的核心传感器模式。语音助手使用麦克风侦听语音并分析信号以识别关键字，这些关键字可以作为简短命令或互联网搜索的触发器。扬声器检测和波束形成功能让会议系统能够提高音质，降低噪音；声学或加速度计数据等时间序列特征可以捕捉机器磨损和即将发生故障的迹象，以便触发预防性维护。

本地时间序列处理和分析对于需要人机交互 (HMI) 或工业监控的嵌入式应用大有裨益。机器学习和 AI 对于信号分析和发现数据中难以察觉的模式非常有效。本文将介绍两个独立 Sitara Arm 微处理器 (AM62x 和 AM62Ax) 上的两个关键字检测应用，进而展示其高性能和低延时。

### 用于关键字检测的时间序列处理和分析

时间序列信号有多种不同的算法和用例。关键字检测和命令识别对于 HMI 应用来说是很重要的功能。在许多应用中：

- 使用集成 DSP 进行传统信号处理 (傅里叶变换、重采样、数字滤波器)。
- 先使用 CPU/MCU 边缘的小型机器学习模型 (隐马尔可夫模型、神经网络)，然后再将信号数据发送到功能强大的云服务器，云服务器可以使用更大、更全面且更准确的模型进行语音识别和分析。随着嵌入式处理器的改进，本地计算功能减少了对云服务器上辅助模型的需求，但产生的代价是重复的资金成本、延时增加以及隐私泄露。

关键字检测任务的时间序列分析需要几个预处理步骤。这些步骤通常包括重采样、滤波、加窗、傅里叶变换等技术 [1]，以及梅尔频谱图和倒谱 (又名 MFCC [2])。除了传统的信号处理算法和独立分量分析 (ICA) 等统计算法之外，神经网络也越来越多地用于特定领域的任务，例如关键字检测、语音转录、语音生成、说话人识别和异常检测。这些领域中的神经网络利用包含 CNN、RNN 和变换器在内的各种技术。

### 用于时间序列分析的微处理器

AMx 系列工业微处理器 (MPU) 能够使用 CPU 资源进行时间序列分析以及各种不同任务。AM62x 和 AM62A 等处理器最多包含四个频率为 1.4GHz 的 Arm® Cortex® A53 CPU。针对更密集的应用和算法，DDR 的内存容量和速度都已经增强 (相比更受限的片上 SRAM)，这对于语音转录等音频分析任务至关重要。关键字检测小字典的 CPU 利用率很低，因此可以和其他任务一起运行。

TI 的 Linux Processor SDK 和 MCU+ SDK 提供了许多软件工具和驱动程序来加速评估和开发。对于这些 SoC，Linux 是最方便且可扩展性很强的操作系统。Debian (从 SDK v9.0 开始) 使用 “apt” 框架简化附加软件包 (不属于基本 SDK) 的安装过程，从而简化了在某些 SoC 上的开发。在 Debian 上，用于时间序列信号处理的软件包可以通过 apt 安装，添加到 Yocto 构建中，从 Ubuntu 等主机进行交叉编译，或直接在目标上构建；特定于 Python 和 Node.JS 等编程语言的库也可以通过各自的打包框架安装在目标上。

该 SDK 包括多个开源机器学习运行时环境，例如 `onnxruntime` 和 `tensorflow-lite`，支持多种神经网络类型（包括 CNN 和 RNN），用于分析音频和其他类型的时间序列信号（例如加速度计数据）。

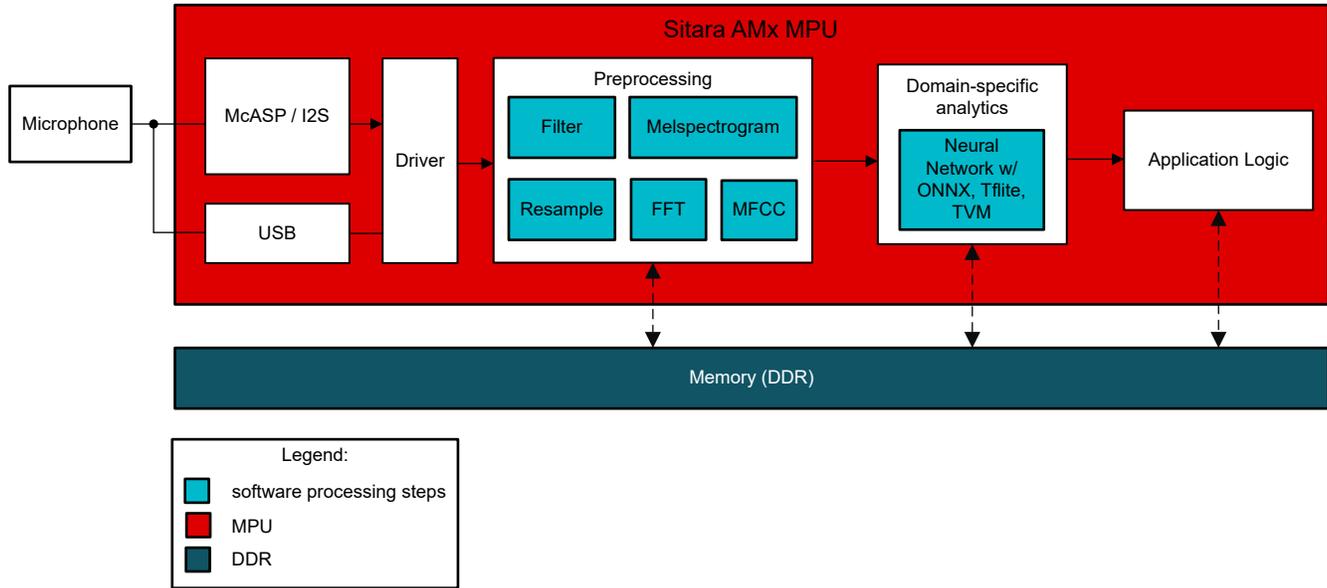


图 1. 在 AMx MPU 上使用神经网络的关键字检测应用的音频处理流程方框图

## 关键字检测应用

Github 上有两个演示应用 [3]，演示了基于 CPU 的关键字检测。这两者都是在 Linux Processor SDK 上使用 Python3 开发的，并利用 python 库进行音频采样、预处理和运行预训练的神经网络。语音数据通过 USB2.0 麦克风输入到 SoC 中。一个演示使用 `matchboxnet` [4] 在一个包含 35 个单词的词汇表中识别命令；另一个演示使用 `MLCommons` 的关键字识别模型 [5] 在一个包含 12 个单词的词汇表中进行语音识别。这两个演示在梅尔频域中使用类似的预处理技术。这两个应用的性能如表 1 所示。

表 1. 16kHz 下 1 秒音频的处理速度，包括使用卷积神经网络进行预处理和关键字检测

器件	运行时规格	预处理时间 (MFCC)	网络 1 : Matchboxnet 推理时间 (四核)	网络 2 : ML Commons Tiny KWS 推理时间 (单核)
AM62A	DDR : 3733MT/s, 32 位 LPDDR4 CPU: 1.25GHz <sup>(1)</sup> 4x A53	38ms	8ms	2.2ms
AM62x	DDR : 1600MT/s, 16 位 DDR4 CPU: 1.25GHz 4x A53	40ms	20ms	2.8ms

(1) 使用 0.75V 内核电压时，支持的最高频率为 1.25GHz。0.75V 内核电压可提供差异化的低功耗能力，比最高 1.4GHz 速度所需的 0.85V 电压更为标准。

这两个应用的预处理步骤包括将音频流从 48kHz 重采样到 16kHz 并计算 MFCC。MFCC 的计算方法是使用短期傅里叶变换、对信号进行平方、使用梅尔滤波器组进行滤波、计算对数以及计算离散余弦变换。在 python3 中，这些步骤是使用 `numpy` 和 `librosa` 库的组合来执行的。

通过 `librosa` 库提供的多核并行化对于音频预处理很有帮助。然而，像此处所述的小型神经网络不一定会从并行处理中获得改善：`matchboxnet` 在 4 核上具有更好的性能，而 `tinyML` 在单核上具有最佳性能。请注意，这些应用的目的是概念验证，并且这些实现可以被认为更接近可实现的最低性能，而不是最高性能。

在这两个模型中，预处理时间占主导地位，因为这些都是小型神经网络，只需要在很小的词汇表中确定用户说出了哪个单词。随着神经网络规模和计算要求的增加，内存开始成为瓶颈。这就是为什么 AM62A 在 `Matchboxnet` 上的性能明显优于 AM62x，原因就在于 AM62A 的 DDR 能力是 AM62x 的 4 倍以上，而在 `MLCommons` 网络上的差异则没那么明显。无论如何，这两个任务都可以轻松地与其他任务一起实时运行，因为这两个应用需要的 CPU 资源不超过总体 CPU 资源的 5%。

## 结语

关键字检测等时间序列分析在各种应用中仍然是很有用的部分，尤其是对于那些具有 HMI 要素的应用。在许多情况下，通用 CPU 资源足以完成时序信号处理，允许此类应用与其他任务（包括其他时序分析）并行运行。对于 AM62x 和 AM62A 处理器，由于其高性能和低延时（1 秒数据的延时小于 60ms），因此可以在 CPU 内核上执行时间序列分析和音频分析，并保留大量资源给其他任务。

## 参考文献

1. [https://en.wikipedia.org/wiki/Fourier\\_transform](https://en.wikipedia.org/wiki/Fourier_transform)
2. [https://en.wikipedia.org/wiki/Mel-frequency\\_cepstrum](https://en.wikipedia.org/wiki/Mel-frequency_cepstrum)
3. <https://github.com/TexasInstruments/edgeai-keyword-spotting>
4. [https://catalog.ngc.nvidia.com/orgs/nvidia/teams/nemo/models/commandrecognition\\_en\\_matchboxnet3x2x64\\_v2](https://catalog.ngc.nvidia.com/orgs/nvidia/teams/nemo/models/commandrecognition_en_matchboxnet3x2x64_v2)
5. [https://github.com/mlcommons/tiny/tree/master/benchmark/training/keyword\\_spotting](https://github.com/mlcommons/tiny/tree/master/benchmark/training/keyword_spotting)

## 重要声明和免责声明

TI“按原样”提供技术和可靠性数据（包括数据表）、设计资源（包括参考设计）、应用或其他设计建议、网络工具、安全信息和其他资源，不保证没有瑕疵且不做任何明示或暗示的担保，包括但不限于对适销性、某特定用途方面的适用性或不侵犯任何第三方知识产权的暗示担保。

这些资源可供使用 TI 产品进行设计的熟练开发人员使用。您将自行承担以下全部责任：(1) 针对您的应用选择合适的 TI 产品，(2) 设计、验证并测试您的应用，(3) 确保您的应用满足相应标准以及任何其他功能安全、信息安全、监管或其他要求。

这些资源如有变更，恕不另行通知。TI 授权您仅可将这些资源用于研发本资源所述的 TI 产品的应用。严禁对这些资源进行其他复制或展示。您无权使用任何其他 TI 知识产权或任何第三方知识产权。您应全额赔偿因在这些资源的使用中对 TI 及其代表造成的任何索赔、损害、成本、损失和债务，TI 对此概不负责。

TI 提供的产品受 [TI 的销售条款](#) 或 [ti.com](#) 上其他适用条款/TI 产品随附的其他适用条款的约束。TI 提供这些资源并不会扩展或以其他方式更改 TI 针对 TI 产品发布的适用的担保或担保免责声明。

TI 反对并拒绝您可能提出的任何其他或不同的条款。

邮寄地址：Texas Instruments, Post Office Box 655303, Dallas, Texas 75265

Copyright © 2023，德州仪器 (TI) 公司