

RISC-V助推AloT芯片 设计与应用

时擎科技 于欣





CONTENTS

目录

1

AloT芯片市场 趋势与需求 2

为什么RISC-V 适合AloT芯片 3

时擎TIMES-FORMER技术

4

时擎AT系列 芯片及应用

AloT芯片市场趋势及需求

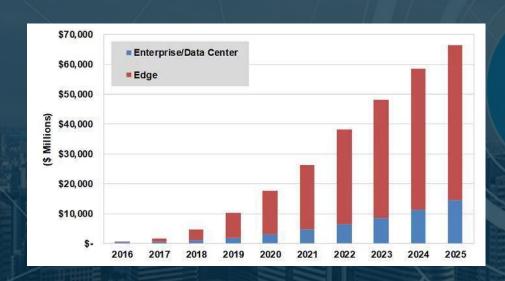
终端AloT芯片市场快速增长,AloT芯片典型需求主要包括算力、低功耗、低成本和灵活性

AloT即Al+loT,指在物联网设备中加入人工智能处理能力。通过本地处理数据缩短系统反应时间,节省海量数据传输带宽,提高隐私性和安全性,从而提高设备附加值和投资回报率。当前AloT应用体现出智能化、碎片化、功耗成本敏感等特点。

AloT芯片典型需求

- 百GOPS级别的处理能力 覆盖大部分应用场景。
- 低功耗的要求依然迫切 能耗比成为关键因素。
- 成本依然敏感 语音芯片0.5~1美金,视频芯片 2~3美金。
- 通用性、灵活性需求 满足碎片化的市场需求、算法演进需求。

全球终端AI芯片市场趋势



数据来源: Tractica

为什么RISC-V架构适合AloT芯片?



RISC-V架构在技术上符合需求、商业上成本优势明显、开源开放使国产设计自主可控

技术层面

RISC-V指令集精简、模块化、可扩展的架构优势带来低成本、低功耗和灵活等特点, 完美符合边AloT芯片需求。

商业层面

RISC-V无指令集架构授权费、精简的指令集使得相同设计能获得更小的芯片面积从而带来巨大成本优势。

其他层面

RISC-V指令集架构开源、开放, 任何人都可免费试用,不受美 国出口管制,符合我国对处理 器、芯片自主可控的需求。

RISC-V是一种开放、免费、可扩展的计算机指令集,2010年由伯克利大学提出。2015年RISC-V基金会成立后得到迅速的发展和推广,目前已有超过300家公司和组织会员。RISC-V是一个最新的、清晰、简约、开源的指令集.它以过去指令集设计所犯过的错误为鉴,目标是让它在从最小的到最快的所有计算设备上都能有效工作。



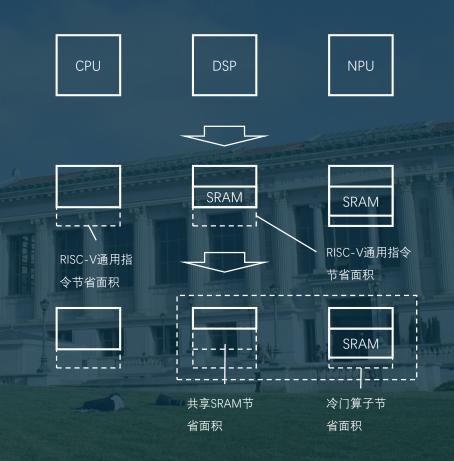
RISC-V架构释放AloT芯片设计潜能

RISC-V架构的开放性打开了AloT芯片CPU+DSP+NPU的计算架构优化的大门

• CPU+DSP+NPU在大部分边缘智能芯片中面积、功耗占比较高。e.g. 30%~50%

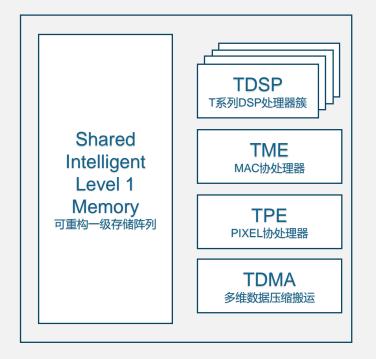
• RISC-V的精简、模块化属性为处理器" 瘦身",大幅提高单位面积算力。

• RISC-V的可扩展属性释放DSA领域专用架构创新潜能,进一步提高能效比。



TIMESFORMER边缘智能计算架构

一种适应未来3-5年边缘智能计算需求的可重构、可伸缩计算架构





可重构

根据计算任务动态调整协处理器 算子、DSP处理器簇的配置、数 据通路和存储结构,重构出最适 合该任务的计算部件组合及相应 的数据流。



高能效

精密的加速算子设计、数据流优化兼顾灵活性与高能效比。 40nm工艺实测1.5TOPS/W, 28nm工艺预计>3TOPS/W。



可伸缩

仅需通过更改编译配置即可生成 包含不同算子、不同算力大小、 适应不同应用场景的代码, 大幅 加速芯片开发周期、降低开发成 本。



高性能

各类主流神经网络平均计算效率接近 90%。超长指令字+宽向量架构的多核 TDSP处理器结合可复用的MAC、 PIXEL协处理器提供强大的DSP处理能 力。



TIMESFORMER主要核心部件

所有处理器、IP均为自主设计且算力均可配置

可重构RISC-V协处理器



- TME可重构MAC协处理器
- 支持人工智能算法中的各类MAC密集运算
- 支持视觉算法中的各类MAC密集运算



- TPE可重构PIXEL协处理器
- 支持人工智能算法中池化、逐点、激活等运算
- 持视觉算法中的常见像素操作、滤波

T系列RISC-V DSP处理器



- T100 MCU/DSP处理器, 4级流水
- RISC-V IMCF+Packed-SIMD扩展
- 针对语音处理



- T200 DSP处理器, 7级流水
- RISC-V IMCF + 向量扩展
- 针对语音、视觉处理



- T300 DSP处理器, 11级流水
- RISC-V IMCF+向量扩展+VLIW扩展
- 针对视觉处理、SLAM



TIMESFORMER可伸缩架构覆盖各类AloT边缘智能计算场景

仅需更改配置即可生成针对不同边缘智能计算应用场景的硬件组合





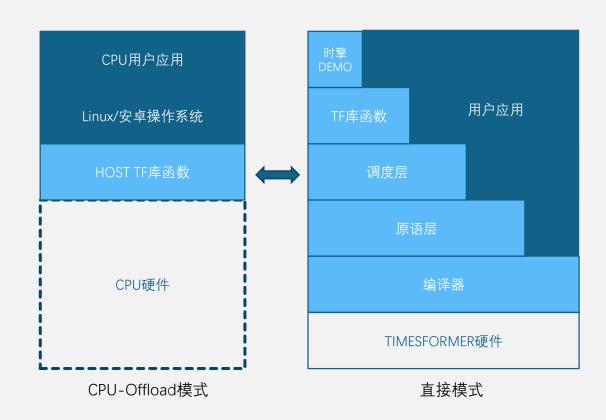


主要应用场景	语音识别、本地NLP	人脸/行为识别、物体检 测、多模态交互	多目标检测、图像语义、 vSLAM、多模态交互
AI算力	30GOPS-300GOPS	300GOPS-1TOPS	1TOPS-8TOPS
DSP核类型	T100	T100/T200	T300
DSP核数量	1/2/4/8	1/2/4/8/16	1/2/4/8
DSP核向量宽度	32位	32位/64位/128位	64位-512位
1级存储大小	32KB-128KB	64KB-256KB	128KB-2MB

TIMESFORMER配套软件栈: TFSDK

TFSDK提供丰富的编程接口、库函数和辅助工具,加速用户应用开发、移植、差异化定制

- CPU-Offload模式可快速移植用户已有的在主流应用处理器上的代码。 HOST TF库函数支持对接OPENCV、 Tensorflow-Lite的Linux、安卓版本。
- 直接模式用户在TIMESFORMER上开发应用,TFSDK提供了从DEMO参考设计到底层编程的各层接口,满足客户差异化定制需求。
- 时擎DEMO包括语音识别、人脸识别、 SLAM等。TF库函数包括人工智能算 法库、类OPENCV算法库以及语音 DSP算法库。



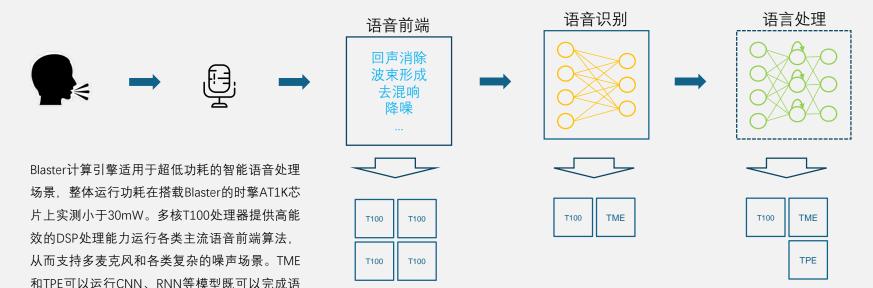


TIMESFORMER高能效智能处理应用案例: 语音识别

搭载Blaster计算引擎的AT1K芯片以超低功耗运行语音前端及语音识别算法

音识别功能如关键词识别, 也可以运行一些轻量

级的自然语言处理任务如简单的问答。

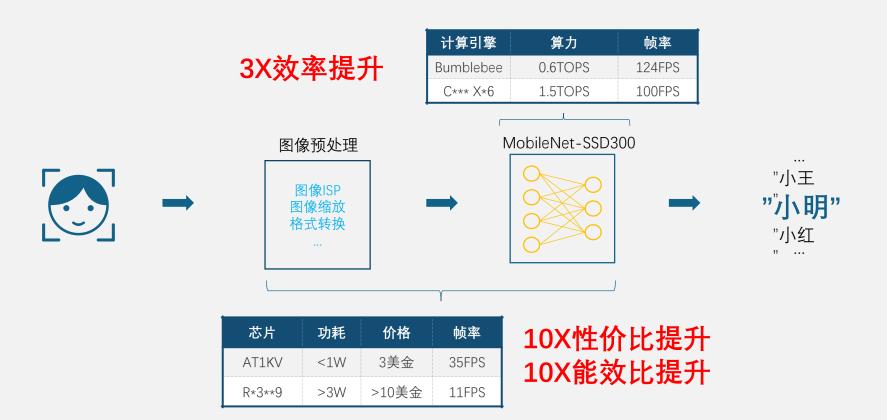


处理功耗低至30mW



TIMESFORMER高能效智能处理应用案例: 人脸识别

搭载Bumblebee计算引擎的AT1KV芯片以更高的计算效率、更低的功耗和成本达到更高的人脸识别帧率

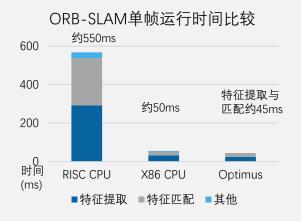




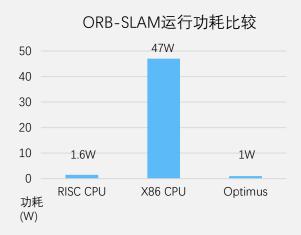
TIMESFORMER高能效智能处理应用案例: 视觉SLAM

Optimus计算引擎以极高的能效比满足视觉SLAM实时计算需求

满足实时计算需求



大幅节省计算功耗



适用广泛的低功耗场景









瓦级视觉SLAM实时处理



时擎基于RISC-V架构的AloT芯片应用场景举例

时擎基于RISC-V架构的AloT方案相较传统方案具备低成本、高能效比、编程友好、灵活适应算法演进等优势

应用场景

通用智能语音 (AT1K)



家庭智能监控 (AT1KV)



智能门禁/门锁 (AT1KV)



支持算法

- AEC/BF/去混响/降噪(Blaster)
- 关键词、句识别(Blaster)
- 本地NLP、简单对话(Blaster)

- 多目标检测、识别 (Bumblebee)
- 图像预处理(ISP、Bumblebee)
- 图像编码(H.264/JPEG)
- 关键词、简单对话(Bumblebee)

- 人脸检测识别(Bumblebee)
- 图像预处理(ISP、 Bumblebee)
- 关键词、简单对话(Bumblebee)
- 指纹、声纹检测(Bumblebee)

系统需求

- 100条以上指令词、句
- · 快速人声唤醒(硬件VAD)
- · 低功耗待机(硬件VAD)
- 接口: 麦克风阵列、扬声器、无线连接

- 多目标实时检测、识别
- 1080P@30FPS录像
- 用户隐私保护
- 接口:摄像头、红外传感器、 无线连接、麦克风、扬声器

- 实时检测(>30FPS), >100张人 脸
- 快速唤醒(<1s), uW级待机
- 接口:摄像头、显示屏、指纹 传感器、红外传感器、麦克风 、扬声器



THANKS!

