Intelligent Vision for the Industrial, Automotive and IoT Edge with the i.MX 8M Plus Applications Processor

Ali Osman Ors
Director, AI ML Strategy and Technologies
NXP Semiconductors
Introduction to NXP applications processors
Introduction to i.MX 8M Plus SoC
i.MX 8M Plus machine vision features
  • ISP
  • NPU
AI ML and machine vision software enablement
Scalable Compute Platforms

TRADITIONAL KINETIS AND LPC MCUs

i.MX RT CROSSOVER MCUS

CORTEX-M CORES

HIGH-PERFORMANCE

SECURE, ENERGY EFFICIENT

ML ACCELERATION

3D & 2D GRAPHICS

ALWAYS LISTENING

Embedded Flash

I/Os

Wi-Fi

BLE

Zigbee

Thread

Wi-Fi

BLE

Zigbee

Thread

TSN end nodes

TSN switch

BROAD CONNECTIVITY PORTFOLIO

FUNCTIONAL INTEGRATION

© 2022 NXP Semiconductors
i.MX Series of Applications Processors

Future i.MX 9 Families
Automotive, Industrial & IoT
Edge Processing

i.MX 8 Family
Advanced Graphics, Vision & Performance

i.MX 8M Family
General Purpose
Edge Processing

i.MX 8ULP Family
Ultra Low Power with Graphics

i.MX 8X Family
Safety Certifiable & Efficient Performance
Arm® v8-A (32-bit/64-bit)

i.MX 7 Family
Flexible Efficient Connectivity

i.MX 7ULP Family
Ultra Low Power with Graphics
Arm® v7-A (32-bit)

© 2022 NXP Semiconductors
i.MX Series of Applications Processors

i.MX 6
- 12 product families
- Offers software and pin-pin compatibility
- Arm® v7-A

i.MX 8 Family
Advanced Graphics, Vision & Performance

i.MX 8M Family
General Purpose Edge Processing

i.MX 8ULP Family
Ultra Low Power with Graphics

i.MX 8X Family
Safety Certifiable & Efficient Performance
Arm® v8-A (32-bit/64-bit)

i.MX 7 Family
Flexible Efficient Connectivity

i.MX 7ULP Family
Ultra Low Power with Graphics
Arm® v7-A (32-bit)

i.MX 8M Plus
2.3 TOPS NPU, Vision (ISP), 1080p video encode/decode, advanced HMI, multi-display, USB3/2xGbE/PCIe

i.MX 8M
4K HDR, 4K video decode, advanced HMI, multi-display USB3/GbE/2xPCIe

i.MX 8M Mini
Versatile multimedia applications processor with 1080p video encode/decode, HMI graphics acceleration, USB/GbE/PCIe

i.MX 8M Nano
Entry-level multimedia applications processor with HMI graphics acceleration, USB/GbE
# i.MX 8M Plus Applications Processor with ML & Vision Engines

## i.MX 8M PLUS BLOCK DIAGRAM

<table>
<thead>
<tr>
<th>Security</th>
<th>Main CPU Platform</th>
<th>Display</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm® TrustZone®</td>
<td>2 x/4 x Arm Cortex®-A53</td>
<td>HDMI 2.0a Tx (eARC) with PHY</td>
</tr>
<tr>
<td>DRM Ciphers</td>
<td>32 KB I-cache</td>
<td>MIPI-DSI (4-lane) with PHY</td>
</tr>
<tr>
<td>Secure Clock</td>
<td>32 KB D-cache</td>
<td>1 x LVDS Tx (4 or 8-lane) with PHY</td>
</tr>
<tr>
<td>eFuse Key Storage</td>
<td>Arm Neon™</td>
<td></td>
</tr>
<tr>
<td>Random Number</td>
<td>FPU</td>
<td></td>
</tr>
<tr>
<td>32 KB Secure RAM</td>
<td>512 KB L2 Cache (ECC)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Secondary Cores</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cadence® Tensilica® HiFi4 DSP</td>
<td>SP/DIF Tx and Rx</td>
</tr>
<tr>
<td></td>
<td>Cortex-M7</td>
<td>eARC (HDMI)</td>
</tr>
<tr>
<td></td>
<td>868 KB On-Chip RAM</td>
<td>ASRC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8-ch. PDM Microphone Input</td>
</tr>
<tr>
<td>System Control</td>
<td>Machine Learning</td>
<td>Connectivity and I/O</td>
</tr>
<tr>
<td>Smart DMA x3</td>
<td>Neural Processing Unit: 2.3 TOPS</td>
<td>2 x USB 3.0/2.0 Dual-Role with PHY</td>
</tr>
<tr>
<td>XTLAL</td>
<td>Graphics</td>
<td>2 x GbE Ethernet with IEEE® 1588, AVB</td>
</tr>
<tr>
<td>PLLs</td>
<td>3D Graphics: GC7000UL</td>
<td>(One also supports TSN)</td>
</tr>
<tr>
<td>Watchdog x 3</td>
<td>2D Graphics: GC520L</td>
<td>2x CAN/CAN FD</td>
</tr>
<tr>
<td>PWM x 4</td>
<td>Video</td>
<td>1 x PCIe® Gen 3 – 1-lane</td>
</tr>
<tr>
<td>Timer x 6</td>
<td>1080p60 H.265, H.264, VP9, VP8 decoder</td>
<td>L1 Substates</td>
</tr>
<tr>
<td>Secure JTAG</td>
<td>1080p60 H.265, H.264 encoder</td>
<td>4 x UART 5 Mbit/s, 6 x I2C, 3 x SPI</td>
</tr>
<tr>
<td>Temperature Sensor</td>
<td>Vision</td>
<td>External Memory</td>
</tr>
<tr>
<td></td>
<td>Dual Camera ISP (2 x HD/1 x 12MP) HDR, dewarp</td>
<td>32-bit LPDDR4/DDR4 (Inline ECC)</td>
</tr>
<tr>
<td></td>
<td>2 x MIPI-CSI (4-lane) with PHY</td>
<td>3 x SDIO3.0/MMC5.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dual-ch. QuadSPI (XIP) or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 x Octa/SPI (XIP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NAND Controller (BCH62)</td>
</tr>
</tbody>
</table>
i.MX 8M Plus Applications Processor with ML & Vision Engines

4x Cortex® A53 @ 1.8GHz
- Speech recognition
- Object detection
- Gesture recognition

HiFi4 DSP @ 800MHz
- Voice/Keyword recognition
- Speech enhancement
- Noise reduction

GC7000UL 3D Graphics Engine (GPU) @ 1GHz
- Object detection classification
- Floating point support

Cortex-M7 @ 800MHz
- Keyword detection
- Sensor fusion
- Anomaly detection

Neural Processing Unit (NPU) @ 1GHz
- Multi-camera classification and detection

Two-channel Image Signal Processor (ISP)
- De-warping and image enhancement
### i.MX 8M Plus Target Applications

#### ML and Industrial Automation
- Machine vision and robot controller
- Industrial computer, gateways, HMI
- Printers and scanners
- Machine visual inspection
- Factory automation

#### Smart Home, Building & City
- Safety, security and surveillance
- Fleet analytics
- Traffic monitor and flow optimization
- Vision payment systems
- Targeted advertisement
- Service drones
- Alarm and AI server hubs
- Home patient and elderly monitor

#### Consumer Audio/Voice Systems
- Surround sound and sound bars
- Audio/video receiver
- Immersive audio products
- Wireless or networked smart speakers
- Personal assistant
- Voice-assisted products
ISP (Image Signal Processor)
• **Image Signal Processor (ISP) basic function**
  • Converts the image color code from raw Bayer (output of the image sensor) to YUV so it can be processed by the SoC
  • Provides additional processing to improve the image quality:
    • HDR extracts maximum image detail in high contrast scenes
    • De-Warp: Fisheye lens or low-cost lens geometry correction
    • Image Enhancement

• **i.MX 8M Plus - ISP Benefits:**
  • Low latency and high performance
  • Lower BOM cost for vision system
  • Higher product longevity
Bayer de-mosaicing and filtering (including denoising, sharpening and blurring)
Defect pixel cluster correction (DPCC)
Color processor (CPROC)
Chromatic aberration correction (CAC)
Denoise
Histogram
Lens shading correction (LSC)
Wide Dynamic Range (local tone mapping)
Color noise removal (CNR)
Automatic white balance measurements (AWB)
Exposure measurement for AE (AEC/AGC)
Auto focus measurement (AF)
2-exposure and 3-exposure DoL/Staggered HDR
AL/ML
NPU (Neural Processing Unit)
i.MX 8M Plus NPU Subsystem

Programmable Engine Unit
- 128-bit vector processing. INT 8/16/32b, FLOAT 16/32b.
- Most flexible programming unit

Vision Engine
- Provides advanced image processing functions

Universal Storage Cache
- Local memory and L1 cache to pass data amongst NPU modules

Tensor Processing Core
(3 instances)
- INT 8/16b, FLOAT 16b
- Non-convolution layers.
- Multi-lane processing for data shuffling, normalization, pooling/unpooling, LUT, etc.
- Network pruning support, zero skipping, compression

Neural Network Core
(6 instances)
- 2.3 TOPs INT8
- Convolution Layers, RELU, Max Pooling, and Compute Bounded Fully Connected Layers
- Scalable 8- and 16-bit integer multiply-accumulate (MAC) engine for tensor operations
- Specialized NN hardware supports activation and pooling
- Supports variety of NN topologies:
  - Convolutional (CNN): MobileNet, YOLO, etc
  - Recurrent (RNN, GRU, LSTM): Deep Speech 2, etc.
NPU Performance Increase for Quantized Models

- Measured on i.MX 8M Plus with 1 Cortex A53 CPU core vs. NPU normalized to 1 GHz for both cores
- Both A53 and NPU are 8-bit quantized
Machine Vision
Software Enablement
eIQ® ML Software Development Environment

- **eIQ Toolkit**
  - eIQ Marketplace
  - Value-Add Solutions and Services
  - Bring Your Own Data Workflow
    - Data Curation → Model Selection → Model Training, Optimization, Quantization → Model Validation
  - Bring Your Own Model Workflow
    - PyTorch → ONNX

- **Base SW enablement**
  - eIQ Inference
    - TensorFlow Lite
    - ONNX Runtime
    - DeepViewRT
  - eIQ Pipelines
    - AUDIO/VOICE
    - VISION
    - ANOMALY DETECTION

- **Application SW Packs**
  - State Monitor
  - Person Detection
  - Face recognition
  - Occupancy Detection
  - Complex Diagnosis
  - ....

© 2022 NXP Semiconductors
- **Gstreamer** offers an open-source industry-standard solution to handling media components for embedded devices.

- **NNStreamer** provides a set of GStreamer plugins allowing developers to apply neural networks, attach related frameworks (including ROS, IIO, FlatBuffers, and Protocol Buffers), and manipulate tensor data streams in GStreamer pipelines easily and execute such pipelines efficiently.

Gstreamer: https://gstreamer.freedesktop.org/
NNStreamer: https://nnstreamer.ai/
Machine Vision Use Case
Available Demos

Multimedia
- GStreamer
  - Video Test Source
  - Camera Preview
  - Camera using VPU
  - Multi-Cam Preview
- ISP
  - ISP Control Demo
  - Video Dump Demo
- Audio
  - Audio Record
  - Audio Play

GPU
- GLES2
  - Vivante Launcher
  - Cover Flow
  - Vivante Tutorial
  - Bloom
  - Blur
  - DF Graphics Basic 2D
  - Eight Layer Blend
  - Fractal Shader
  - Line Builder 101
  - Model Loader
  - S03 Transform
  - S04 Projection
  - S06 Texturing
    - Mapping
    - Mapping Refraction
  - OpenVG 2D
    - Tiger G2D

Machine Learning
- NNStreamer
  - Object Classification
  - Object Detection
  - Pose Detection
  - Brand Detection
- PyelIQ
  - Object Classification
  - Object Detection
  - Mask Detection

Not all demos listed are available on all boards
NNstreamer (ML Vision) Demos

- Object classification demo
- Object detection demo
- Pose detection demo
- Brand classification demo
ISP Demos

• The “ISP control” demos allow users to control the following:
  • De-warp
  • Frames per second
  • Auto white balance
  • Color processing
  • De-mosaicing
  • Gamma control
  • Filtering
  • Black level subtraction

• New demo in 2022 Q1: users can use GUI to dump unprocessed camera data onto a USB drive
Conclusion

The i.MX platforms are some of the most versatile applications processors families for multimedia and display applications

• The i.MX 8M Plus SoC offers
  • Camera capture, network and display drivers / frameworks
  • Multiple ML frameworks available, with multiple inference HW supported (NPU, GPU, CPU, DSP)
  • Multimedia accelerators (2D GPUs, 2.5D GPUs, 3D GPUs...) and industry-standard APIs
  • eIQ ML SW Development Environment
  • Reference code and demos targeting machine vision applications
Embedding Intelligence at the Edge
- www.nxp.com/ai

i.MX 8M Plus product page:
- www.nxp.com/imx8mplus

eIQ® ML Software Development Environment
- www.nxp.com/eiq

eIQ ML/AI Training Series
- www.nxp.com/mltraining

NXP Application SW Packs
- www.nxp.com/appswpack
Thank You