DEVELOP SMART COMPUTER VISION SOLUTIONS FASTER

WITH INTEL® COMPUTER VISION SDK & OTHER ADVANCED SOFTWARE TOOLS

Hosted by the Embedded Vision Alliance
Presented by Intel Corporation
AGENDA

1. Trends Driving a Need for Computer Vision
2. Computer Vision & Deep Learning – Value Together
3. Optimize Your Applications with the Right Tools
   - Intel® System Studio
   - Intel® Computer Vision SDK
   - Intel® Media SDK
   - Intel® SDK for OpenCL™ Applications

Goals
- Show how integrating computer vision can bring smart capabilities to great solutions
- Provide a technical introduction to each so you can get started
VIDEO: THE NEW FRONTIER

Multiple sources: IHS, Markets & Markets, Strategy Analytics, Intel research

- IP Cameras
- Central Management, Archive and Analytics
- Drones
- Video Gateways
- Servers & Recorders
- Security & Surveillance
- Traffic Control
- Smart Home
- Manufacturing
- Retail
- Healthcare
- Infrastructure
Developers need tools that...

- Are comprehensive and easy to use
- Quickly help resolve defects in complex systems
- Offer insight into sources of excess power consumption
- Enable & accelerate performance - demanding & unique, competitive use cases

...and take full advantage of Intel hardware accelerators
DEEP LEARNING BREAKTHROUGH

Image Recognition

- 2010
- Present

Error
- 0%
- 8%
- 15%
- 23%
- 30%

Using Deep Learning

Human

97% person
END-TO-END DISTRIBUTED INTELLIGENCE

SMART CAMERAS

VIDEO GATEWAYS

DATA CENTER / CLOUD

ACCELERATE & DIFFERENTIATE WITH INTEL SOFTWARE TOOLS
END-TO-END INTELLIGENCE

27.2 Mbps for 1080p/25FPS/H.264

END-TO-END INTEL® ARCHITECTURE (SMART CAMERA)

Faster output, lower BPS

License Plate Detection

OCR

Comparison

27.2 Mbps if no license plate detected

0 Mbps if no license plate detected

0 Mbps if no license plate detected
END-TO-END ARTIFICIAL INTELLIGENCE FOR AUTOMATED DRIVING

**VEHICLE**
- Driving Functions
- Environment Modeling
- Sensor Fusion
- Anomaly Detection

**NETWORK**
- Captured Sensor Data
- Real-time Model, SW, FW Updates
- Data Formatting, Storage, Management, Traceability

**CLOUD**
- Model Training
- Model Inference
- Compress Model
- Universal Models
- Reasoning Systems

5G

ML DL RB LB

OpenVX and the OpenVX logo are trademarks of the Khronos Group Inc.
OpenCL and the OpenCL logo are trademarks of Apple Inc. used by permission by Khronos
INTEL HARDWARE IS HETEROGENEOUS

Skylake
6th Generation Core

Kabylake
7th Generation Core

Apollo Lake

Intel® Pentium
• J4205, J4200
Intel® Celeron
• J3455, J3355
• N3350, N3450
Intel Atom
• E3950, E3940, E3930
UNLOCK HARDWARE CAPABILITIES

NEW!

INTEGRATE VISUAL UNDERSTANDING
Intel® Computer Vision SDK Beta
Linux*/Yocto* version available

HETEROGENEOUS CUSTOM DEVELOPMENT
Intel® SDK for OpenCL™
Application Development

ACCELERATE VIDEO PROCESSING
Intel® Media SDK
for Embedded Linux®, Windows® & Open Source
Intel® C++ Compilers deliver impressive performance on embedded applications powered by Intel Atom processors

Coremark Pro* benchmarks running on Intel Atom processors

### 32-bit mode

<table>
<thead>
<tr>
<th>Compiler</th>
<th>GCC 6.1.0</th>
<th>Intel® C++ 16.0</th>
<th>Intel® C++ 17.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clang 3.8</td>
<td>1.11</td>
<td>1.25</td>
<td></td>
</tr>
</tbody>
</table>

### 64-bit mode

<table>
<thead>
<tr>
<th>Compiler</th>
<th>GCC 6.1.0</th>
<th>Intel® C++ 16.0</th>
<th>Intel® C++ 17.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clang 3.8</td>
<td>1.09</td>
<td>1.11</td>
<td></td>
</tr>
</tbody>
</table>


Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark & MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information & performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. Benchmark Source: Intel Corporation - Optimization Notice: Intel’s compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets & other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User & Reference Guides for more information regarding the specific instruction sets covered by this notice. Notice revision #20170304.
YOUR BUILDING BLOCKS FOR IMAGE, SIGNAL & DATA PROCESSING APPS
Intel® Integrated Performance Primitives (Intel® IPP)

What is Intel® IPP?
Provides developers with ready-to-use, processor-optimized functions to accelerate Image, Signal, Data Processing & Cryptography computation tasks

Why use Intel® IPP?
• High Performance
• Easy to use API’s
• Faster Time To Market
• Production Ready

How to get Intel® IPP
Intel System Studio
Intel Parallel Studio XE
Free Tools Program

Optimized for

Supports

Addressed

Image Processing
• Medical Imaging
• Computer Vision
• Digital Surveillance
• Biometric Identification
• Automated Sorting
• ADAS
• Visual Search

Signal Processing
• Games (sophisticated audio content or effects)
• Echo cancellation
• Telecommunications
• Energy

Data Compression & Cryptography
• Data centers
• Enterprise data Managements
• ID verification
• Smart Cards/wallets
• Electronic Signature
• Informationsecurity/cybersecurity

Find out more at: software.intel.com/intel-ipp

Optimization Notice
Copyright © 2017, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.
What is Intel® TBB?
A highly templated C++ library designed to simplify adding parallelism to your application by taking advantage of all the CPU's on a single device or across multiple devices (heterogeneity).

Why use Intel® TBB?
• High Performance
• Easy to use API's
• Faster Time To Market
• Production Ready

Optimized for

Supports

Addresses

How to get Intel® TBB
Intel System Studio
Intel Parallel Studio XE
Free Tools Program
Open Source site

Applications
• Artificial Intelligence & Automation
• Image processing
• Any solution needing sophisticated threading

Find out more at: software.intel.com/intel-tbb
Contact us: software.intel.com/forums/intel-threading-building-blocks
Get Faster Code Faster with Accurate Data & Meaningful Analysis

- Accurate CPU, GPU & threading data
- Memory access & storage analysis
- Powerful data analysis & filtering
- Data displayed on the source code
- Easy set-up, no special compiles

“Last week, Intel® VTune™ Amplifier helped us find almost 3X performance improvement. This week it helped us improve the performance another 3X.”

Claire Cates
Principal Developer
SAS Institute Inc.

Learn More: intel.ly/vtune-amplifier-xe
A NEW SDK FOR COMPUTER VISION

INTEL® COMPUTER VISION SDK BETA

MACHINE LEARNING INFERENCE

Model Optimizer
- > OpenVX code generation
- > IR
- > Inference Engine SDK
Same interface for CPU, GPU, ...

Many Topologies: Lenet, Alexnet, Googlenet, VGG, SSD + YOLO (soon)

OPTIMIZED LIBRARIES

OpenVX™
OpenCV

UTILITIES

Vision Algorithm Designer

Prototype

Traditional Computer Vision

OpenCL implementations
Model Optimizer Bridges Train and Deploy

- Generate [OpenVX code](#)
- Generate [Intermediate Representation (IR)](#)
- Optimize the network
  - Node fusion
  - Node merging
  - Batch normalization
- Calculate and dump the normalized and converted weights/biases ([normalization factor](#) can be supplied by user if learning phase is skipped)
$ cd /opt/intel/computer_vision_sdk_2017.0.090
$ source bin/setupvars.sh
$ cd mo/bin

$ export FRAMEWORK_HOME=
/home/user/Desktop/MO_LAB/caffe/build/lib/

$ ./ModelOptimizer --target APLK -i \n -d /home/user/Downloads/caffe-master/models/bvlc_reference_caffenet/deploy.proto.txt \n -w /home/user/Downloads/caffe-master/models/bvlc_reference_caffenet/bvlc_reference_caffenet.caffemodel \n -f 1 \n -p FP16 \n -o artifacts

Start working...
Framework plugin: CAFFE
Target type: APLK
Network type: CLASSIFICATION
Batch size: 8
Precision: FP16
Layer fusion: true
Output directory: artifacts
Custom kernels directory:
Network input normalization: 1
Writing binary data to: artifacts/CaffeNet/CaffeNet.bin

Optimization Notice
Copyright © 2017, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.
MODEL OPTIMIZER GENERATED OPENVX CODE

1 Train
2 Prepare model
3 Inference
SETUP CODE

// Plugin name (MKLDNNPlugin | CLDNNPlugin | OVXPlugin)
string pluginName = "clDNNPlugin";
string pluginPath = "path/to/plugin/"+pluginName;

// Load plugin
InferenceEnginePluginPtr plugin(pluginPath);
InferenceEngine::InferenceEnginePluginPtr _plugin(plugin);

// Read network
InferenceEngine::CNNNetReader network;
network.ReadNetwork(pathToModel + ".xml");
network.ReadWeights(pathToModel + ".bin");

InputsDataMap inputs = network.getNetwork().getInputs();

// Allocate input and output blobs here

// Load model to plugin
sts = _plugin->LoadNetwork(network.getNetwork(), &dsc);
Mat frame, frame2;
for (;;) {
    cap >> frame;

    // resize to expected size (in IR .xml)
    resize(frame, frame2, Size(227, 227));

    // run inference
    long unsigned int framesize = frame2.rows*frame2.step1();
    ConvertImageToInput(frame2.data, framesize, *input);

    sts = _plugin->Infer(*input, *output, &dsc);

    // check status here...

    // get top classifier label
    int blobsize = output->size();
    float *data = output->data();
    float max = 0;
    int maxidx = 0;
    for (int i1 = 0; i1 < blobsize; i1++) {
        if (data[i1] > max) {
            max = data[i1];
            maxidx = i1;
        }
    }

    // do something with classification
    imshow( "frame", frame2 );
    if (waitKey(30) >= 0) break;
}
OPENCV VS. OPENVX

- Industry standard
  - 47K user community
  - >14M downloads
- Community driven open source
- >2500 algorithms
- CPU C++, growing list of OpenCL/CUDA implementations
- Standard scheduling, no automatic tiling across functions, etc.

- Emerging standard
- Created for power optimized heterogeneous HW development
- Vendor driven, all or partial closed source
- ~50 algorithms
- Designed for fixed function, may be implemented in C++, OpenCL, etc.
- Automatic graph level optimizations (tiling, etc.)

OpenVX and the OpenVX logo are trademarks of the Khronos Group Inc.
OpenCL and the OpenCL logo are trademarks of Apple Inc. used by permission by Khronos
**DEVELOPMENT FLOW OPTIONS**

**Vision Algorithm Designer**
- Build Pipelines
- Debug
- Performance feedback

**OpenVX™ C/C++ API**
- Use with familiar IDEs
- Interoperable with other libraries, SDKs & programming models

```c
vx_context context = vxCreateContext();
vx_image input = vxCreateImage( context, 640, 480, VX_DF_IMAGE_U8 );
vx_image output = vxCreateImage( context, 640, 480, VX_DF_IMAGE_U8 );

vx_graph graph = vxCreateGraph( context );
vx_image intermediate = vxCreateVirtualImage( graph, 640, 480, VX_DF_IMAGE_U8 );
vx_node F1 = vxF1Node( graph, input, intermediate );
vx_node F2 = vxF2Node( graph, intermediate, output );

vxVerifyGraph( graph );
vxProcessGraph( graph ); // run in a loop
```
**THEORY OF OPERATION:**
**INTEL® MEDIA SDK / INTEL® MEDIA SERVER STUDIO**

**Media accelerator framework**
Codec based
High level/parameter interface
3 operations

Good option for:
- Accelerated video encode, decode
- (and short list of frame processing)

More Information
- Media Server Studio
- Media SDK
- Intel Media Code Samples

**Out of scope:**
* audio, containers, networking...
CODECS + FRAME PROCESSING USE FIXED FUNCTION + EUS

**Video Encoding**
ENC = EU+VDBox VME (MB type, motion vectors, bit budget/BRC)
PAK = VDBox (residue packing & entropy coding)
VDENC = low power encode (6th Generation Core® & forward)

**Video Decoding**
BSD = VDBox decode

**Media Fixed Function**

- **VDBOX**
  - EU
  - EU
  - EU
  - EU
  - Sampler
  - Caches

- **VEBOX**
  - EU
  - EU
  - EU
  - EU
  - EU
  - Sampler
  - Sampler

**VPP**
Video Processing Hardware Acceleration Layer

**VPHal**
- **VEBox**
  - Deinterlacing
  - Denoise (Luma/Chroma)
  - Frame Rate Conversion
  - Color space conversions
  - Composition/alpha blending
  - Scaling
Basic Decode Flow

**Initialize**

DecodeFrameAsync (bitstream in)

- MFX_ERR_MORE_SURFACE
- MFX_ERR_MORE_DATA

Main loop

More input

Finish

(MFX_ERR_MORE_DATA indicates all surfaces drained)

Input finished

Drain loop

DecodeFrameAsync (null in)

- MFX_ERR_MORE_SURFACE
- MFX_ERR_MORE_DATA

**Expected Return Codes for DecodeFrameAsync**

- **MFX_ERR_MORE_SURFACE**
  - A new surface is required to proceed – this is where decode will write its output

- **MFX_ERR_MORE_DATA**
  - More input bitstream data is required to proceed

- **MFX_WRN_DEVICE_BUSY**
  - Hardware device is unable to respond. This is an expected output for normal operation and should clear after a short wait. However, if this state persists more than a few milliseconds this may indicate a problem.

- **MFX_WRN_VIDEO_PARAM_CHANGED**
  - The SDK decoder parsed a new sequence header. Decoding can continue with existing frame buffers. The application can optionally retrieve new video parameters by calling MFXVideoDECODE_GetVideoParam.

- **Other**
  - Other error codes may be bugs. Please contact an Intel support representative for more info.

Optimization Notice
Copyright © 2017, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.

OpenVX and the OpenVX logo are trademarks of the Khronos Group Inc.
OpenCL and the OpenCL logo are trademarks of Apple Inc. used by permission by Khronos
do {
    if (still_reading_file) { // main loop
        sts = mfxDEC.DecodeFrameAsync(&mfxBS, pmfxSurfaces[nIndex], &pmfxOutSurface, &syncp);
    } else { // drain loop
        sts = mfxDEC.DecodeFrameAsync(NULL, pmfxSurfaces[nIndex], &pmfxOutSurface, &syncp);
        if (sts==MFX_ERR_MORE_DATA) break;
    }
    switch (sts) {
        case MFX_WRN_DEVICE_BUSY:
            MSDK_SLEEP(1); // Wait if device is busy, then repeat
            break;
        case MFX_ERR_MORE_SURFACE:
            nIndex = GetFreeSurfaceIndex(pmfxSurfaces, numSurfaces); // Find free frame surface
            break;
        case MFX_ERR_MORE_DATA:
            readsts = ReadBitStreamData(&mfxBS, fSource); // Read more data
            if (readsts!=MFX_ERR_NONE) still_reading_file=0;
            break;
    }
    if (MFX_ERR_NONE!=sts) continue;
    sts = session.SyncOperation(syncp, 60000); // Wait until decode finished
}

// frame data can be used by application now

Add VPP resize and classification code here
INTEL® SDK FOR OPENCL™ APPLICATIONS

SDK Tools

• Kernel analyzer
• Kernel debugger
• Offline compiler
• IDE integration

Implementations for many Intel Hardware platforms

Today’s Focus

GPU
CPU
FPGA
Future

Download from
software.intel.com/intel-opencl
EXTENSIONS MAP

Device-side
   VME

Block IO

VEBox

Image from buffer
mipmap
depth images
packed/planar YUV

Subgroups

Intel Subgroups
Khronos Subgroups
Short Subgroups

Accelerator

Sharing

VAAPI
DX9
DX11
OpenGL

Other

SPIR
Diagnostics
fp16/fp64
atomics
...

For more info: software.intel.com/articles/opencl-intel-graphics-extensions
SHARING APIs IN ACTION

Interop example code (in this tutorial)
Intel® Media SDK/Intel® Media Server Studio samples (sample_multi_transcode, sample_encode)
Hardware Video Motion Estimation

VME is part of the Media Sampler

- Programmable through EUs
- Operates on 16x16 macroblocks
- 1 per sub-slice
  - 2 sub-units (co-issuable)

- Implements key motion estimation operations
  - Inter Motion Estimation
  - Sub-pixel refinement
  - Intra Prediction
  - Many more...

- Programmable general purpose operations

- Optimized for memory bandwidth
- Provides configurable raw compute
- Smarts in the hands of the programmer
INTRODUCING VEBOX

A configurable pipeline of common video processing operations

Before

After

Adaptive Contrast Enhancement (ACE)

Total Color Correction (TCC)
Input Source
- File source
- V4L2 Camera
- Network camera

Video Processing (VPP)
- Resize
- Compose
- Color Conversion

Encode
- H264 HW Encode
- HEVC HW Encode
- MJPEG HW Encode

NEW!
INTEGRATE VISUAL UNDERSTANDING
Intel® Computer Vision SDK Beta

OpenVX/CNN
SSD CNN OpenVX Graph

HETEROGENEOUS CUSTOM DEVELOPMENT
Intel® SDK for OpenCL™ Application Development

ACCELERATE VIDEO PROCESSING
Intel® Media SDK

Optimization Notice
Copyright © 2017, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.

OpenVX and the OpenVX logo are trademarks of the Khronos Group Inc.
OpenCL and the OpenCL logo are trademarks of Apple Inc. used by permission by Khronos
MORE RESOURCES – DOWNLOAD SOFTWARE TO GET STARTED

ACCELERATE VIDEO PROCESSING
Intel® Media SDK
Free Download > software.intel.com/media-sdk

INTEGRATE VISUAL UNDERSTANDING
NEW! Intel® Computer Vision SDK Beta
FREE Download > software.intel.com/computer-vision-sdk

CUSTOMIZE WITH OPENCL
Intel® SDK for OpenCL™ Applications
FREE Download > software.intel.com/intel-opencl

CPU + system optimization
Specialized hardware acceleration

Free Trial > software.intel.com/system-studio

Optimization Notice
Copyright © 2017, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.
Empowering Product Creators to Harness Embedded Vision

The Embedded Vision Alliance (www.Embedded-Vision.com) is a partnership of 60+ leading embedded vision technology and services suppliers

Mission: Inspire and empower product creators to incorporate visual intelligence into their products

The Alliance provides low-cost, high-quality technical educational resources for product developers

Register for updates at www.Embedded-Vision.com

The Alliance enables vision technology providers to grow their businesses through leads, ecosystem partnerships, and insights

For membership, email us: membership@Embedded-Vision.com
Intel technologies’ features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

Cost reduction scenarios described are intended as examples of how a given Intel®-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Intel, the Intel logo, the Intel Inside logo, Intel Atom, Quark and Mashery are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.
Legal Disclaimer & Optimization Notice

INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS”. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Copyright © 2017, Intel Corporation. All rights reserved. Intel, Pentium, Xeon, Xeon Phi, Core, VTune, Cilk, and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

---

Optimization Notice

Intel’s compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804