AI Power FPGA Based Video Transcoding For Real Time Application

Presented By

Aupera
Making Video Alive

Name  Roy Liao
Title   CEO
Date   2018.10.16
Agenda

1. Video Market Outlook
2. Aupera FPGA Based Solution
3. Aupera Development Framework
Video Market Outlook

Video Is Everywhere

- **245 million** Surveillance Cameras
- **$40 billion** Video-on-demand market
- **344 million** Live broadcasting viewers
- **82%** Internet traffic will be video by 2021
- **80%** Of data centre will be video data

Sources: Cisco, IDC, IFSEC, IBM reports
Increasing Computing Complexity

Video Codec Marketshare

- H.264
- H.265
- VP9
- AV1
- H.266
- Other

Source: Xilinx
Fast Changing Networks

ANNN, Madaline, Back Propagation, CNN, DCNN, Floating Point

Perceptron, Belief Net, TDNN, ResNET, 40 Years


Deep Belief Neural Net, ZFNet, ResNet, VGG Net, LeNet, Alex Net, GoogLeNet, DCCN, Transformer Net, Fast RCNN, WaveNet, VDCNN, Fractal Net, 8-bit to 1-bit, Variable Precision

Future

Source: Xilinx
Video Challenges & Aupera Solution

Video Processing Challenges

- Bandwidth cost
- Computing complexity
- Energy cost
- Video content real time analytics
- Variety of consumers end devices

Vertical System Optimization:
- Algorithm → FPGA → HW → SW
- FPGA based high density video transcoding
  - 380 channel 1080p30 in 3U chassis
- FPGA built-in CV/ML engine
  - Real time object detection and classification
- 90% energy saving with FPGA + Arm
Xilinx MPSoC Architecture
Aupera Architecture

Video + AI Applications

Streaming Framework
(FFmpeg, Gstreamer, etc.)

Aupera Acceleration Stack
(xfOpenCV, CHaiDNN, DPU, VPM)

Hardware Platform-AUPV205, AUP2600
Video Transcoding Application

- Operation Center
- Edge Nodes
- OSS
- 3rd Party CDN (option)
- Private RTN
- RTMP/HLS
- RTMP/HLS
- DB Records

Companies and Technologies:
- Aupera
- Xilinx
Video + AI Applications

Scene Understanding

Pedestrian Segmentation

Object Classification and Detection
ROI Enhanced Video

```bash
while true; do gstreamer-1.0 filesrc location=/gstreamer/data/live-004.mp4 ! qtdemux ! h264parse ! omxh264dec ! auperadetector ! queue ! omxh265enc gop-length=30 periodicity-idr=30 control-rate=1 target-bitrate=120 quant-i-frames=40 qp-mode=0 ! h265parse ! queue ! rtph265pay ! udpsink host=10.53.170.117 port=50000 max-lateness=-1 qos-dscp=60 async=false max-bitrate=500000000; done
```

Gstreamer plugin

Video Decoder

AI/CV Application

Video Encoder

H.264

ROI Optimized 120Kbps

120Kbps

ROI Optimized 120Kbps

Aupera

Making Video Alive

Xilinx Developer Forum

Xilinx

Object Detection

Face examples

Off-line Training

Feature Extraction

Classifier

Classification Result

Linear SVM
Latent SVM

Feature vector \((x_1, x_2, \ldots, x_n)\)

FHOG
HOG

Search for faces at different pyramid levels

Non-face examples

Face examples

Non-face examples

Aupera

XILINX DEVCORNER FORUM
FHOG

- Spatial case of histogram of oriented gradients
- Reduces feature size with no loss of information
  - 36 dimensional regular HOG → 31 dimensional FHOG
  - 9 orientations → Contrast insensitive features
  - 18 orientations → Contrast sensitive features
FPGA Acceleration Results

- Implemented on Aupera v205
- FHOG pyramid feature extraction
- 50X Acceleration

<table>
<thead>
<tr>
<th></th>
<th>ARM</th>
<th>FPGA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1080</td>
<td>1.2 s</td>
<td>24 ms</td>
</tr>
<tr>
<td>540</td>
<td>730 ms</td>
<td>15 ms</td>
</tr>
<tr>
<td>270</td>
<td>190 ms</td>
<td>3.8 ms</td>
</tr>
</tbody>
</table>

FHOG

![Graph showing comparison between ARM and FPGA for different image sizes: 1080, 540, and 270. ARM times are 1.2 s, 730 ms, and 190 ms, respectively, while FPGA times are 24 ms, 15 ms, and 3.8 ms, respectively.]
Aupera Development Framework

Closing the gap between ideas and deployment for video+AI applications

Application Templates
- Video Transcoding
- ROI Enhanced Enc
- Video Conference
- Metadata Extraction
- Stream Mixer
- tracking/capturing

Streaming Media Frameworks
- CV
  - Accelerators
  - xfOpenCV
- Codec
  - VPM*/VQ+
  - Hardware AVC / HEVC
- DL
  - CHaiDNN
  - Deephi DPU

Hardware / Software Platform Generator
- PL Design
- Customized Boot/Root FS
- Application Config Tool
- Xilinx SDSoC (Vivado, PetaLinux…)

- Video Focused
- Application Oriented
- Optimized Accelerators
- On top of Xilinx SDSoC

Note*: VPM stands for Aupera Video Processing Module

Aupera Development Framework

Xilinx Existing Environment

Xilinx
Vivado, PetaLinux…”
HW/SW Platform Generator

Accelerate Video Application Development

- **One-Click** to generate runnable images and application templates
- Customize application
- Online evaluation

### Example Platform Generator

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Hardware Model</th>
<th>SDSoc SDK</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aupera Technologies Inc.</td>
<td>V205-A0-Z7EV-1</td>
<td>2018.2p3</td>
<td>VCU, VPM(2 scl+1 pm), Peta default, MGMT agent, VLAN</td>
</tr>
<tr>
<td></td>
<td>V205-A1-Z4EV-1</td>
<td>2018.2p2</td>
<td>VCU, VPM(1 scl), CW(ROI-HOG), Peta default, MGMT agent</td>
</tr>
<tr>
<td></td>
<td>V205-A1-Z7EV-1</td>
<td>2018.2p1</td>
<td>VCU, VPM(1 scl), DL(CNN MobileNet), Peta default, VCU, VPM(1 scl)</td>
</tr>
<tr>
<td></td>
<td>V205-A1-Z7EV-2</td>
<td>2018.1</td>
<td>VCU, VPM(1 scl), DL(DPU), Peta default, MGMT agent, VL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2017.4</td>
<td>VCU, VPM(2 scl+2 mixer+2 pm), Peta default, DHCP, VLAN</td>
</tr>
</tbody>
</table>

### Boot Device

- QSPI-1
- QSPI-2
- eMMC
- TF Card
Video Application Templates

Streaming Mixer

Input Stream → Input Plugin → VCU Decoder → YUV Data → VPM* → Filters → VCU Encoder → Output Plugin → Output Stream

Aupera Development Framework
- Basic Customization
- Advanced Customization

Note*: VPM stands for Aupera Video Processing Module
Aupera HW/SW Platform Highlight

- 380 channels 1080p30 Video streams
- H.264, H.265 Transcoding with AI built-in
- Standard Plugin
- API Compatibility
- Easy Deploy
- Smooth Scale Out
- Remote Upgrade

AUP2603
19 inch 3RU
Aupera HW/SW Platform Highlight

33x Advantage in Perf/watt with ZU7EV

1 x Aup2603 (48 ZU7EV)

Video transcoding + AI analytics

Aupera™ Solution | Intel E5 Solution
---|---
Energy Cost | 10x
Transcoding Cost/feed | 3.3x
Space | 10x

30 x Intel E5 Server
Adaptable.
Intelligent.

Aupera
Making Video Alive