Optimizing Video and Image Recognition

ImageVision® Analytics Platform running on Intel® Xeon® processors

Solutions Summary:

- A literal explosion of visual content as internet consumer video traffic is expected to grow to close to 80 percent in 2019, up from 64 percent in 2014.
- Illustration of the challenges in deploying a scalable Visual Analytics platform including hardware/software, unstructured nature of content, and achieving accuracy.
- An end-to-end solution from ImageVision® and Intel that provides an accelerated and scalable solution for deployment in cloud and enterprise.
- Provide technology solutions for enterprises with objective of extracting meaningful revenue from Visual Content.

The patented, proprietary algorithms of ImageVision deliver the scale, speed and accuracy for video and image recognition expected by global brands and content rich media companies serving their consumers. ImageVision achieves scale and lowest Total Cost of Ownership (TCO) of their software solutions using Intel® Xeon® processors.

- Real-time – delivering SLAs to meet global brands and content rich media companies’ requirements across the Internet.
- Accuracy – 95-99% precision, better than human tagging and moderation.
- Speed – image analysis in <250ms.
- Scale – processing 100M videos and images daily.

“We kept asking ourselves: What would be different if your computer could see like a person?”

- Brian Krzanich, Intel CEO

Market Background and Driving Forces:

With billions of videos and photos shared every day using connected devices such as phones, tablets, and cameras, arguably today’s greatest digital asset is the free and abundant supply of user generated content.

According to the Cisco® Visual Networking Index report, globally, consumer Internet video traffic will be 80 percent of all consumer driven Internet traffic in 2019, up from 64 percent in 2014.

This percentage does not include video exchanged through peer-to-peer (P2P) file sharing. The sum of all forms of video (TV, video on demand [VoD], Internet, and P2P) will be in the range of 80 to 90 percent of global consumer traffic by 2019. The content on the Internet is also increasing in modality as the usage of all types of media, including video, audio, images, and text, continue to weave into the fabric of the internet.

Analyzing this from the enterprise perspective, 80 percent of data is unstructured data located within multimedia content and today is largely absent in organizations’ big data initiatives.

The net result is a literal explosion of visual content that could be leveraged to extract meaningful insights, protect intellectual property, and create new types of services or enhance existing services. Now that we have all of this data, the real challenge is what to do with all of it.
The Complexity and Challenges of Multi-Modal Content

With this explosion of diverse Internet content comes some inherent challenges to unlocking its value. Below is a list of just some of the challenging questions today’s enterprises are trying to answer:

• How do you create an end-to-end hardware/software solution that scales as datasets grow and demand volatility increases?

• How do you reach a high enough classifier accuracy to be useful for the domain of interest?

• Once content is classified, how do you generate meaningful insights or devise ways to monetize the patterns discovered?

• How do you handle the unstructured nature of visual content as web content becomes increasingly multi-modal, including diverse types of video and images?

From the customer perspective, the simple desire is to have an end-to-end scalable solution allowing them to easily spot and analyze trends, patterns, and anomalies that lie within the vast amount of data that they have available to them. They can then exploit these insights for a variety of applications such as targeted advertisements, security/surveillance, content protection, and many more.

An End-to-End Solution for Visual Search

In much the same way the human brain recognizes people, places, and things, ImageVision* has developed evolutionary methods to describe objects that mimic the complex decision making the brain uses to recognize objects. This recognition process involves proprietary and patented shape features technology and is key to improved performance.
ImageVision methodologies separate foreground from background, accounting for shadows and lighting variations, recognizing objects present in a wide range of orientations, describe the visual shape and texture in a repeatable language, and store the visual shape features in a searchable format. Once an object is recognized, ImageVision utilizes mathematical formulas to describe each object visible by a camera into a set of visual shape features that uniquely describes each object and provides for millisecond searches delivering relevant, organized data upon which an organization’s big data strategy can execute on and deliver on that sought after ROI.

Visual Search based on ImageVision's Visual Analytics software solutions provides enterprises around the world video and image analytics for:

- Illicit Video and Image Detection – aka, ImageSPAM
- Face Detection – with Facial Biometrics
- Nudity Detection – Nudity Filter for Visual Content
- Text Detection – OCR for Visual Content
- Altered Images – i.e. detecting edited or manipulated images
- Image Ballistics – similar to bullet ballistics for guns, but for cameras

**Intel® Xeon® and Xeon Phi™ Processors and Deep Learning**

As compute intensive and highly parallelizable applications such as Image and Video search continue to scale, clouds continue to get more and more visual in nature, requiring new architectures and approaches to hardware deployments. A good example of this is the training of deep neural networks (DNNs). Not only do DNNs require a significant amount of parallel compute cycles, but they are also historically the bottleneck to application deployment due to their long time to train. There is also the challenge of deploying these applications at scale in enterprise and cloud environments where the software programming model matters and code portability with performance is not only a desire but a requirement in order to speed time to deployment.

Enter Intel® Xeon Phi™ processors, based on Intel® Many Integrated Core (Intel® MIC) architecture, and the industry-leading performance and energy-efficiency of the Intel® Xeon® processor E5 family. These processors enable dramatic performance gains and ease of deployment in enterprise and cloud environments. You can now achieve optimized performance for even your most highly-parallel computing workloads such as Deep Neural Network training, while maintaining a unified hardware and software environment.

While a majority of applications will continue to achieve maximum performance using Intel Xeon processors, certain highly-parallel applications such as DNN training will benefit dramatically by using Intel Xeon Phi processors. Each processor features many more and smaller cores, many more threads, and wider vector units. The high degree of parallelism compensates for the lower speed of each individual core to deliver higher aggregate performance for highly-parallel code. You can use Intel Xeon processors and Intel Xeon Phi processors together to optimize performance for almost any workload.
Deployment Models and Customer Proof Points

With enterprises around the world having a diverse set of requirements and constraints, there really is no ‘one size fits all’ solution to address market needs. To meet the varying needs of customers, ImageVision offers the following deployment models:

- **Software as a Service** – applicable for growing enterprises needing Visual Analytics to empower their offerings. This option offers a simple Restful API and no infrastructure to manage but resides outside of corporate IT management. Examples of users include Match.com*, MutualMind*, and TangoMe*.

- **On Premise Deployment** – applicable to global, large scale enterprises looking to manage this workload directly in their on-premise, corporate data centers. This has the advantage of falling within corporate IT but also bears the expense of managing the underlying infrastructural resources. Examples of users include Flickr*, Google*, and Yahoo*.

- **Software Development Kit (SDK)** – applicable for those enterprises integrating/embedding ImageVision offerings into larger, multi-functional applications. This method provides the highest level of customization but also requires the highest level of technical expertise including data wrangling, neural network training, and infrastructure management. An example in this case is Samsung*.

### SUMMARY OF DEPLOYMENT MODELS

<table>
<thead>
<tr>
<th></th>
<th>SAAS</th>
<th>ONSITE</th>
<th>SDK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supported Hardware</strong></td>
<td>Intel® Xeon® and Xeon Phi™ processors</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Software Model</strong></td>
<td>Restful API</td>
<td>Restful API</td>
<td>Applications API</td>
</tr>
<tr>
<td><strong>Infrastructure Ownership</strong></td>
<td>ImageVision*</td>
<td>Customer</td>
<td>Customer or IaaS Partner</td>
</tr>
<tr>
<td><strong>Custom Models (Algorithms)</strong></td>
<td>Limited</td>
<td>Limited</td>
<td>High</td>
</tr>
<tr>
<td><strong>Deep Application Integration</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Data Wrangling Required</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Model (Algorithm) Training</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Intel Visual Cloud Computing Platform

Built on the Intel Xeon and Xeon Phi processor families, the Visual Cloud Computing Platform brings together optimized software libraries, APIs, and hardware solutions for Visual Understanding, Video Delivery, and Cloud Graphics workloads.
About Intel Corporation

Intel (NASDAQ: INTC) is a world leader in computing innovation. The company designs and builds the essential technologies that serve as the foundation for the world’s computing devices. Additional information about Intel is available at newsroom.Intel.com and blogs.Intel.com. Intel's Visual Cloud Computing Platform offers optimized software libraries, APIs, and hardware solutions for Visual Analytics, Media Processing, and Remote Graphics workloads.

Contact joseph.s.spisak@intel.com for more details.

About ImageVision*

ImageVision is the world’s recognized leader offering automated, real-time analysis of videos and images, changing the industry standards by pushing the limits on accuracy, speed and volume of visual content analyzed. ImageVision Visual Analytics provides enterprises of all sizes the tools for identifying trends, patterns, and anomalies based on shared videos and images, thus allowing enterprises the abilities to grow revenues and protect organizations' brands.

Contact Intel@ImageVision.com for more details.

For more information on Intel® Xeon® processors, visit www.intel.com/xeon

---

2 http://www.forbes.com/sites/ciocentral/2013/02/04/big-data-big-hype/
3 Correctly classified content at a rate of 95%-99% in repeated tests on publicly available social media sites’ images. Solution consisted of an Intel Dual Intel® Xeon® processor E5-2650v2 servers with 32GB RAM, ImageVision* C++ based Classify Server and Face Detection module with the software stack to build the Classify Server and Face Detection module including, but not limited to, RedHat* 6.x/CentOS* 6.x Operating System, ImageVision Visual Libraries, Caffe* Deep Learning Framework Libraries, OpenCV* 2.4.10, and Apache* thrift.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked “reserved” or “undefined.” Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information. Do not finalize a design with this information.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel make may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked “reserved” or “undefined.” Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting Intel's Web site at www.intel.com.

Copyright © 2015 Intel Corporation. All rights reserved. Intel, Xeon, Phi, and the Intel logo 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.

Printed in USA ECB/09/15 Please Recycle 001US