AVSS2011 demo session: GPU Enabled Smart Video Node

Sadiye Guler
intuVision, Inc.
10 Tower Office Park, Woburn, MA 01801
sadiye@intuvisiontech.com

Jay Silverstein, Ian Pushee, Xiang Ma,
Ashutosh Morde
intuVision, Inc.
{jay,ian,sean,ashu}@intuvisiontech.com

Abstract

This paper presents an All-in-One video analytics system, a compact, multi-channel, real-time, video monitoring, event detection, alarm notification, event recording and browsing solution implemented on low cost hardware, taking advantage of NVIDIA's GPU CUDA platform. An inventive distribution of video object detection and tracking processing chain between the GPUs and the CPU provides maximum efficiency at the lowest cost.

1. Introduction

Today's commodity graphics hardware contains powerful programmable coprocessors i.e., General Purpose Graphic Processing Units (GPUs) designed to independently process pixels in parallel providing great potential for acceleration of computer vision algorithms. Over the last decade GPUs have evolved faster than CPU's and they continue to do so with NVIDIA's new Fermi architecture [1, 8]. GPUs provide a more attractive and affordable alternative to dedicated special purpose hardware used to be used for speeding up computationally intense algorithms in many applications. In better part of the last decade GPUs have attracted a lot of attention from the computer vision community for speeding up parallel operations in video analysis algorithms [2 and references therein]. Pixel based parallel operations of object detection and frame pre-processing stages of video analysis tasks are well suited to GPU platform and several image and video processing algorithms have been implemented on GPU platforms such as [3, 5, 6, 7, 9].
While tapping into the power of underutilized GPU's has been widely investigated by the research community, to best of our knowledge there hasn't been many products that promote GPU based video analytics solutions.

intuVision All-in-One system described in this paper employs video object detection and tracking algorithms implemented on to NVIDIA's CUDA environment and take advantage of the available GPU processing cores in addition to the available CPU capacity in a low cost, small footprint hardware platform to perform multi-channel video analytics with event based recording in real-time.

Features like auto discovery of available IP cameras, robust zero configuration tracking configuration, generic robust object classification models, advance event detection schedules, SMTP alarm notification and event recording make the All-in-One system an affordable and easy-to-deploy video analytics solution for a wide range of applications.

The remainder of this document is organized as follows: after this introduction, in Section 2 we describe the All-in-One Smart Video Node system briefly. Examples of processing results are highlighted and conclusions are presented in Section 3.

2. All-in-One Smart Video Node

In this Section we first discuss GPU enabled video analytics processing present some experimental speed-up results on various GPU platforms and then present the system features and architecture.

2.1. GPU Enabled Video Analytics

Many of the computationally intense operations of video object detection and tracking can be sped up significantly by use of GPUs as shown in Figure 1. Parallel operations such as background model generation, background model

Configuration	Max Processing fps	Morphological Filtering	Full Process Sequence
2 CPUs	3	26.90	283.64
2 CPU's + GPU	15	1.00	18.96

Figure 1: Comparison of processing results for four 1080p HD video streams on a Dual core Xeon CPU with and without the NVIDIA Quadro GPU. Processing speeds are given in milliseconds per frame.

update and new frame to background comparisons, as well as image filtering operations can be performed much faster on GPU's resulting in large savings in processing time. By performing only the inherently non-parallel operations in CPU and utilizing multi-thread processing within the CPU, significantly boosts the overall computational performance of video tracking systems. Figure 1 illustrates that the addition of a high capacity GPU card increases the processing frame rate of four HD channels on a Dual core Xeon processor from 3fps to 15 fps. Figure 2 shows the increasing processing time gains with increased video resolutions for 3 different processes in the video object detection and tracking chain.

2.2. All-in-One Smart Video Node System

Video Resolution	640x360		960x536		1260x720	
	CPU	GPU	CPU	GPU	CPU	GPU
Process-1	15.59	1.72	42.41	5.38	81.38	8.34
Process-2	7.33	3.90	18.45	9.00	11.05	5.58
Process-3	4.66	2.02	51.50	16.59	23.30	7.62

Figure 2: Comparison of three processes at various video resolutions on a Core Duo Xeon CPU with and without the NVIDIA GTX 220 GPU card. Processing speeds are given in milliseconds per frame.

intuVision All-in-One Smart Video Node is a network video analytics device capable of processing multiple video inputs in real-time, detecting tracking, classifying of objects and issuing and emailing alarms when preset event conditions occur as illustrated in Figure 3. The features of the system include:

- Real-time video object detection, tracking and event monitoring for up to four 4-CIF resolution cameras
- Camera auto discovery- no specific code is not needed to ingest video from IP cameras (Axis, Sony, Bosch, Nuvation, etc.)
- Connection URLs are via auto-discovery (Supports UDP, TCP delivery methods)
- RTSP client for ingesting of various video formats including H264, MPEG, and MPEG4
- Thin client for smart video analytics metadata stream
- Event alarm list and event preview
- Recording of detected video events, display of live and recorded video, and querying of events
- SMTP client for sending e-mail when a specified event occurs
- Advanced Event Alarm scheduling for Holidays, blackout dates for event monitoring

2.3. Smart Video Node Architecture

The All-in-One smart video node has a modular architecture (as depicted in Figure 4) with configurable plug-ins and client modules to best fit the underlying application requirements. The underlying real-time video tracking is the framework descibed in [4]. Additionally we developed a strategy to distribute tasks between the CPU and the GPUs for maximum performance efficiency.

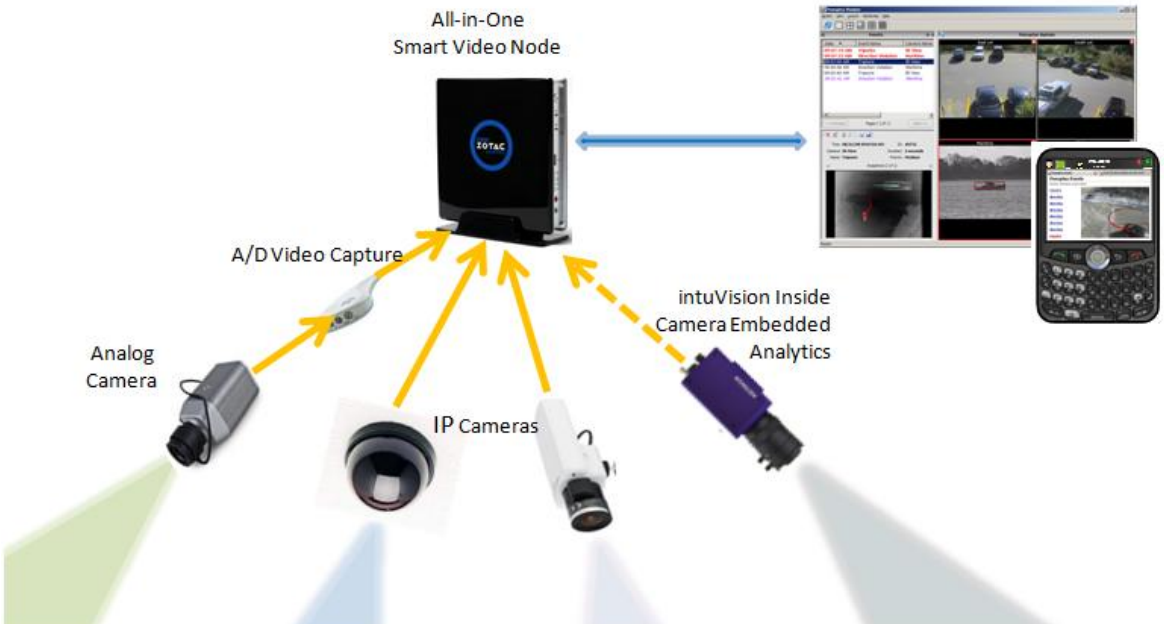

Figure 3: intuVision All-in-One Smart Video Node.

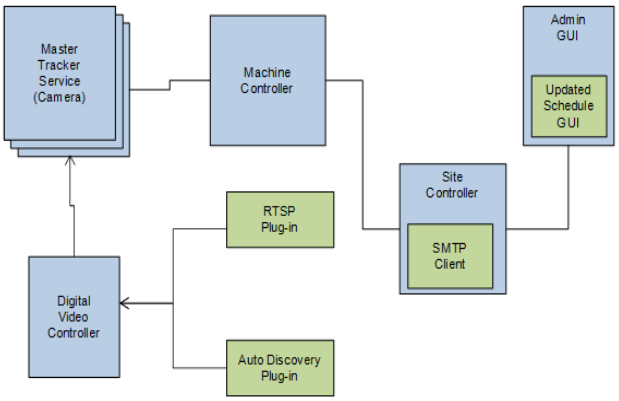

Figure 4: High level architecture of the Smart Video Node.

Master Tracker Service is responsible from managing the operations in the processing chain on CPU or GPU to get the highest overall performance. Our target smart video node platform has a Dual core Intel Atom CPU and NVIDIA ION GPU card with 16 CUDA cores.

2.4. Event Video Recording and Browsing

To make the smart video node an all inclusive video analytics system we added the ability to record and playback video for detected events. These recorded event clips can also be sent around with the event notifications A block flow diagram of video recording capability and components are shown in Figure 5 and the User interface of the all-in-one system with video playback is illustrated in Figure 6.

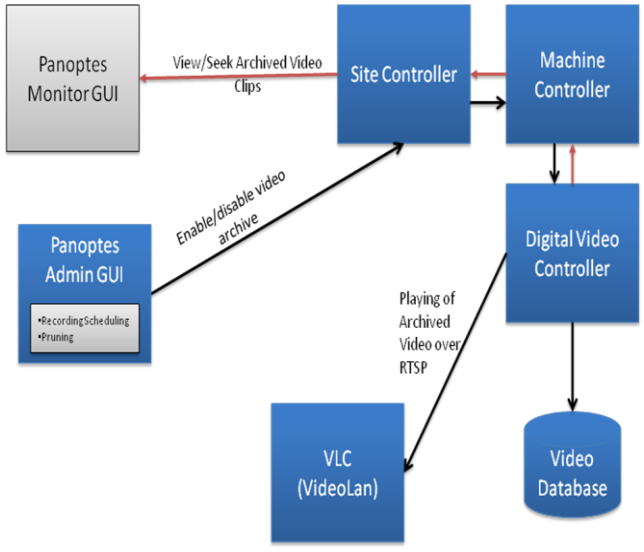

Figure 5: Block Diagram for Smart Video Node video recording.

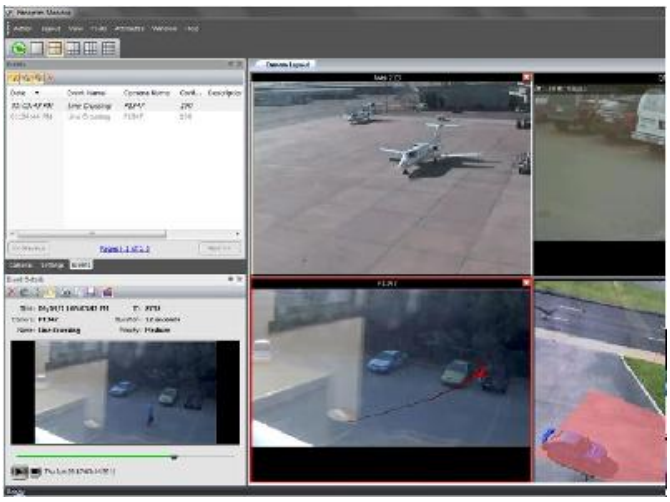

Figure 6: Block Diagram for Smart Video Node video recording.

2.5. Smart Video Node with Camera Analytics

In addition to processing streaming video inputs from various cameras, the smart video node provides a thin client to ingest object detection and tracking metadata streamed from camera embedded analytics and performs event detection on this metadata. A real time streaming server is set up in the smart camera's ARM9 processor streams out the object metadata and frame information for the event detection tasks to be run on the smart video node.

3. Results and Conclusions

We present few important experimental results highlighting the benefits of utilizing GPUs in real time video tracking. Figure 7 on the next page depicts snapshots from the processing of a far field HD video scene with very small tracking targets. The top frame shows a snapshot from the processing of this video without using the GPU while the bottom frame shows a snapshot from the processing with the GPU. Without the GPU the processing frame rate is low and some frames are dropped resulting in missed detections and poor tracking of small targets with low contrast. With GPU enabled processing, several tasks are off loaded to GPU and CPU is freed to handle other tasks making it possible to process each incoming frame. When processed at full frame rate even the very small and low contrast targets are detected and tracked very reliably.

GPU enabled analytics processing makes it possible to perform high quality real-time tracking, on high resolution

video such as this example, or multiple video streams on low cost computers without the need for special-purpose hardware solutions.

In the GPU enabled Smart Video Node system presented here we leverage the parallel stream processing model of the GPUs. Exploiting the high parallelism provided by the GPU architectures several steps of the video tracking algorithms that operate on each pixel independently are performed very fast increasing the overall processing efficiency. As new generation graphics cards such as NVIDIA's Fermi architecture evolve and become widely available our Smart Video Node system is expected to scale to provide even lower cost video analytics per channel.

References

[1] K. A. Bjorke, NVIDIA Corporation,"Image processing using parallel GPU units", Proceedings of SPIE Vol. 6065, January 2006.

[2] M. Gong, A. Langille and M. Gong. "Real-time image processing using graphics hardware: A performance study", International Conference on Image Analysis and Recognition, pp. 1217–1225, 2005.

[3] A. Griesser, S.D. Roeck, A. Neubeck and L.J.V. Gool. "GPU-Based Foreground-Background Segmentation using an Extended Colinearity Criterion". Vision, Modeling, and Visualization (VMV), 2005

[4] S. Guler, J. Silverstein, I. Pushee, X. Ma, A. Morde, Who, What, When, Where, Why and How in Video Analysis: An Application Centric View, Proceedings of AVSS 2010, Boston, August 2010.

[5] P. Labatut, R. Keriven, J.-P. Pons,"A GPU Implementation of Level Set Multiview Stereo". International Conference on Computational Science (4) pp. 212-219, 2006.

[6] L. Mussi, F. Daolio, S. Cagnoni,: GPU-based road sign detection using particle swarm optimization, in IEEE Conf. Intelligent System Design and Applications. pp. 152–157 2009.

[7] S. N. Sinha, Jan-Michael Frahm1, Marc Pollefeys1, Y. Genc, GPU-based Video Feature Tracking And Matching, Technical Report TR 06-012 Department of Computer Science, UNC Chapel Hill, 2006.

[8] The next generation CUDA architecture http://www.nvidia.com/object/fermi_architecture.html

[9] R. Yang, G. Welch, Fast image segmentation and smoothing using commodity graphics hardware. J. Graph. Tools 7(4), pp. 91–100, 2002.

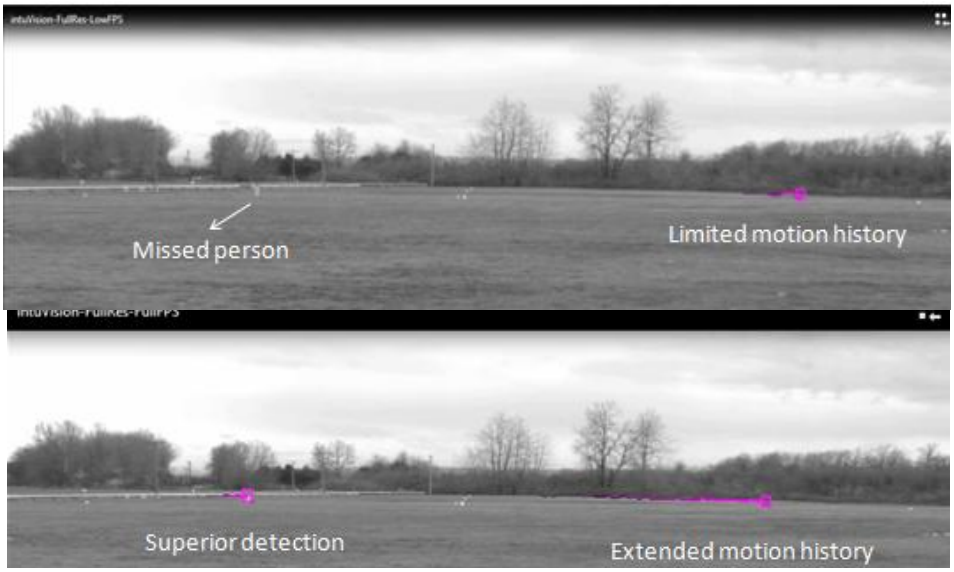

Figure 7: Lower frame rate processing of HD video on a low cost hardware platform misses low contrast small targets (top); GPU enabled full frame rate processing detects and tracks all targets reliably (bottom).