Fermi in Action: Robust Background Subtraction for Real-time Video Analysis

Melvin Wong, Karianto Leman, Nam Trung Pham, Richard Chang, Teck Wee Chua, Gao Feng
Computer Vision & Image Understanding Department
Institute for Infocomm Research (I²R)
{chwong, karianto, ntpham, tewchua, fgao}@i2r.a-star.edu.sg

Overview

Background subtraction is one of the important image processing steps for video surveillance and many computer vision problems such as tracking & recognition. However, robust background subtraction that adapts well to variable environment changes is highly computational and consumed large amount of memory. Thus, its practical application is often limited.

Here, we aimed to expand its usage and tackle vision problems that requires high frame rate camera such as real-time sports analysis, real-time object detection and recognition. Using recent advances in accelerator hardware – NVIDIA Fermi Architecture and taking advantage of heterogeneous computing , we are able to gain good performance that allows to use in these practical applications.

Robust Background Subtraction

The robust background subtraction used in this work is proposed by Liyuan et. al [1]. This is a probabilistic model that builds upon color and gradient information from image sequences. Adapting to time-evolving background and robust against imaging noises, the gradient difference, background difference and temporal difference are defined as:

$$G(x,y,t) = \left\{[\nabla_x F(x,y,t) - \nabla_x B(x,y,t)]^2 + [\nabla_y F(x,y,t) - \nabla_y B(x,y,t)]^2\right\}^{\frac{1}{2}}$$

$$B(x,y,t) = \frac{1}{M^2}\left|\sum_{u=x-M/2}^{x+M/2}\sum_{v=y-M/2}^{y+M/2}[F(u,v,t) - B(u,v,t)]\right|$$

$$T(x,y,t) = \frac{1}{M^2}\left|\sum_{u=x-M/2}^{x+M/2}\sum_{v=y-M/2}^{y+M/2}[F(u,v,t) - F(u,v,t-1)]\right|$$

where ∇_x and ∇_y are the directional gradients by the Sobel operator, *M* is the local average window size, *B(x,y,t)* is the background reference model at time instant *t*, *F(u,v,t)* is the video frame at time instant *t*.

Once the color & gradient information are extracted, binary classifiers are applied to extract background from foreground based on current probability models. The updated information is then fused and refined to extract foreground blobs. Probability and background models are updated based on these new background and foreground information.

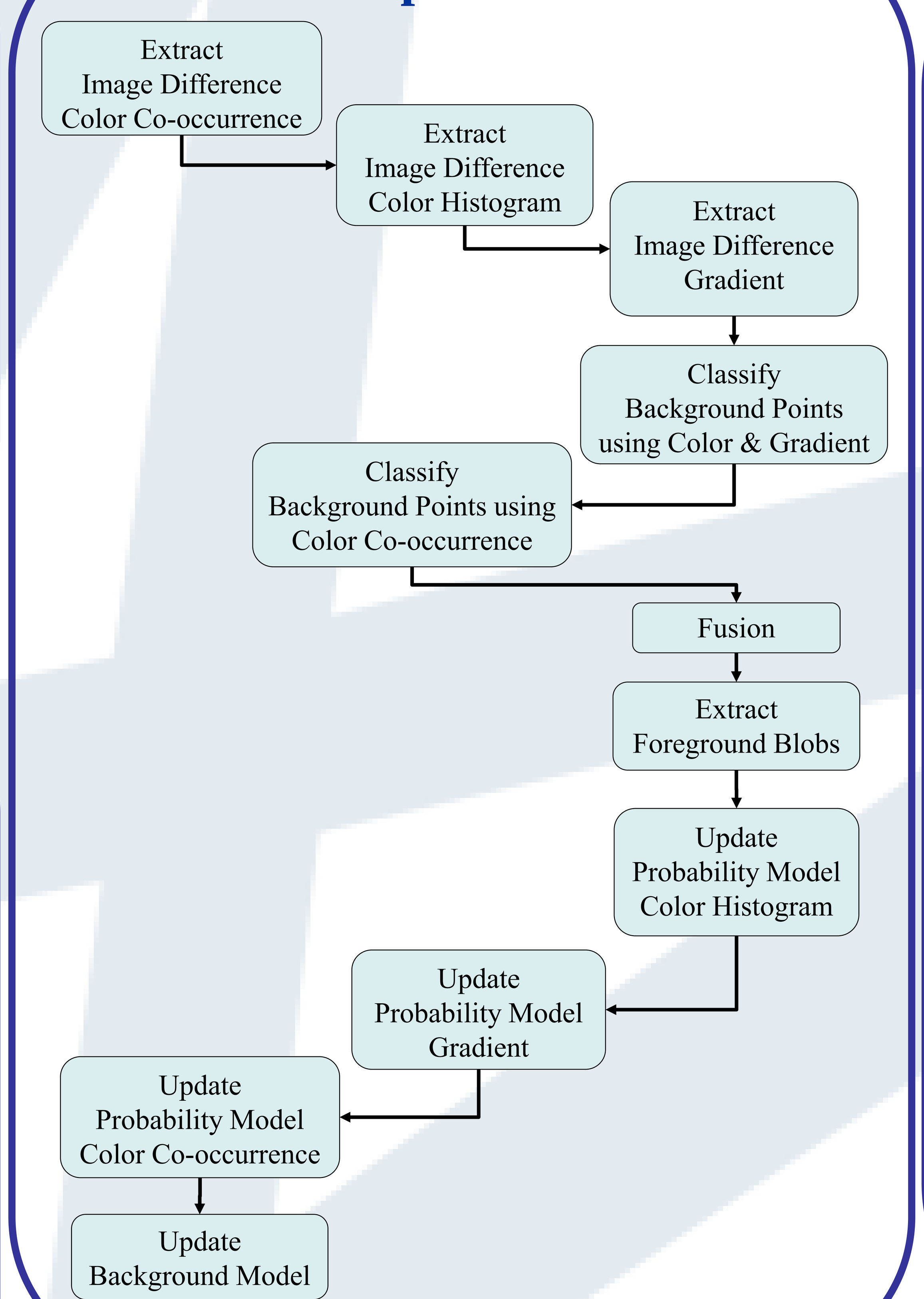
Multi-Core CPU versus Heterogeneous Computing

As the robust background subtraction is highly computational expensive and memory exhaustive, large amount of systems resources is needed for real-time processing in high frame rate camera. This limits the number of higher level vision tasks that can be explored in practical applications.

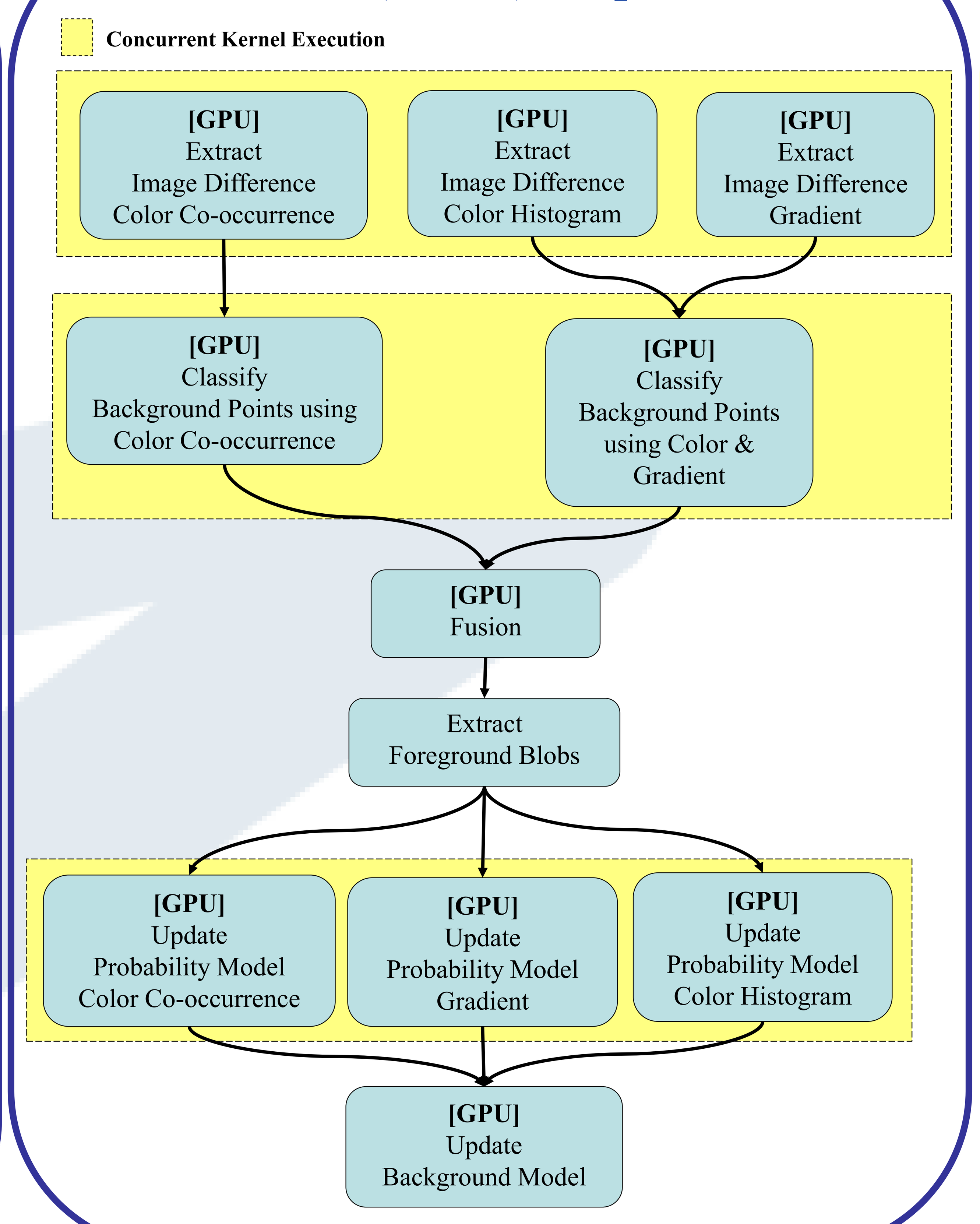
In order to resolve this issue, we migrate highly computational tasks in robust background subtraction to process in GPUs. In this way, CPU load is greatly lessen to allow more higher level vision tasks to be applied practically.

Using the latest NVIDIA Fermi Architecture, we are able to run multiple kernels concurrently with appropriate balance between memory usage per kernel and concurrent threads execution. Thus, enjoy good performance gain and applicable for practical applications that needed high frame rate cameras.

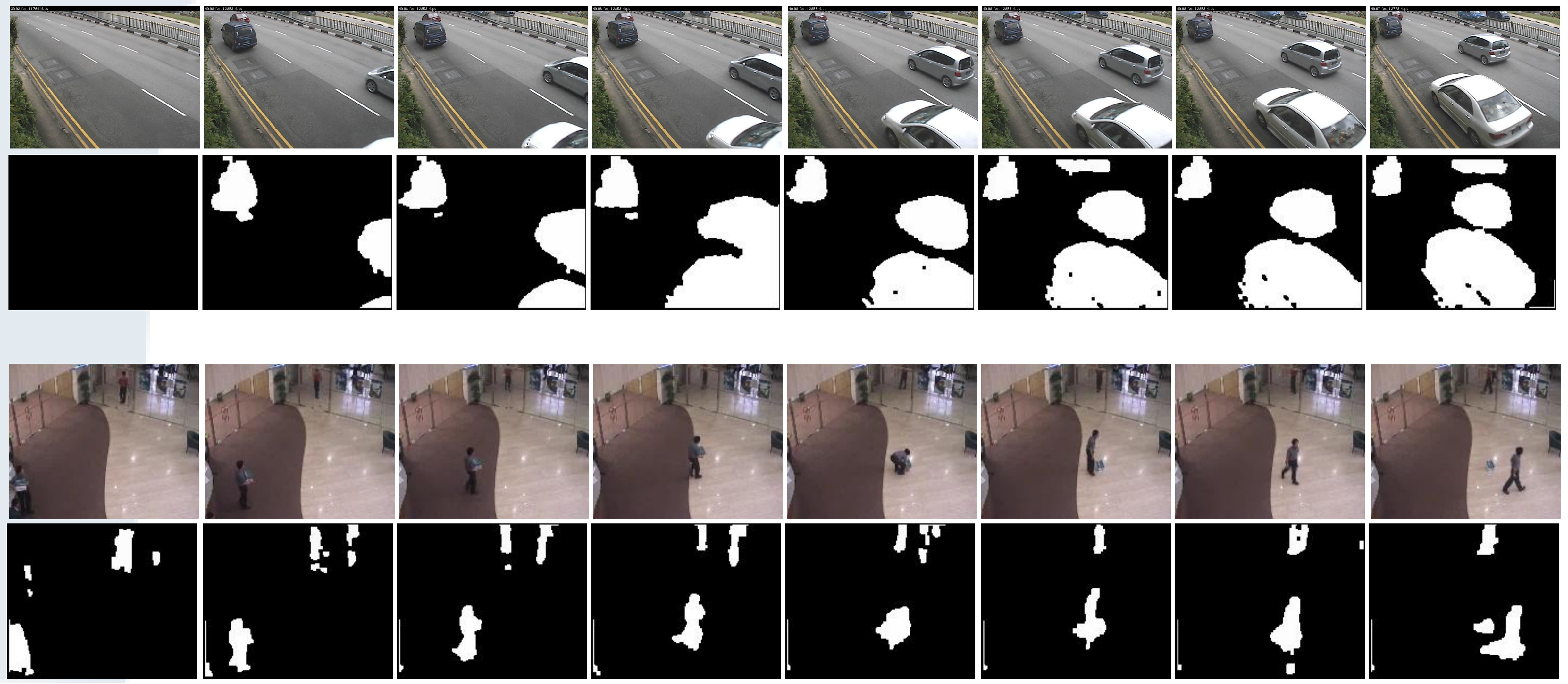
CPU Implementation



CPU + GPU (Fermi) Implementation



Preliminary Experiment Results



Benchmark

Hardware				
Manufacturer	Intel	Intel	Intel	NVIDIA
Model	i3	Xeon 5160	Xeon X5570	470
# cores used	4	8	8	448
Implementation	C/C++	C/C++	C/C++	CUDA
Year	2010	2007	2009	2010
Performance / Cost				
Full System Cost (approx.)	$1K	$6K	$10K	$1.4K
Relative Speedup	1x	1.5x	2.2x	3x
Relative Perf./$	1x	0.25x	0.22x	2.14x

Ongoing Process

1. Scaling up in large scale video surveillance system
2. Migrating whole robust background subtraction algorithm into GPU to minimize slow data latency.
3. Dynamic load balancing of GPU kernels
4. Dynamic task scheduling of robust background subtraction as a service.

Reference

[1] Liyuan L., Ruijiang L., Weimin H., Karianto L., Wei Yun Y., "Adaptive Background Subtraction with Multiple Feedbacks for Video Surveillance", Intl' Symposium on Visual Computing (ISVC), 2005, pp. 380-387.

I²R
A*STAR