Moving Object Segmentation for Security and Surveillance Applications

Goksel Dedeoglu, Ph.D.

Darnell J. Moore, Ph.D.

Vision R&D, Texas Instruments, Inc.

Texas Instruments

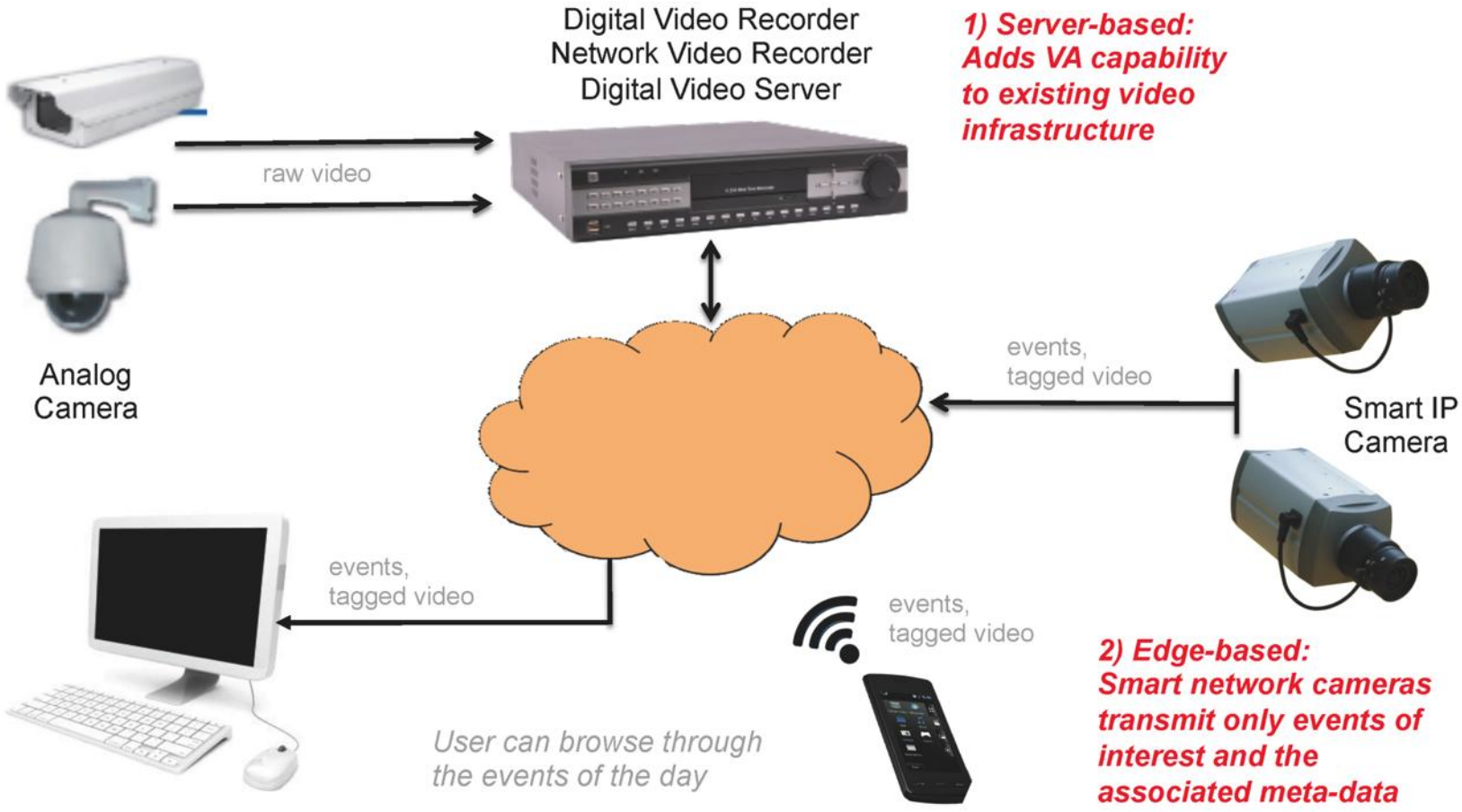
Motivation

- Cameras are everywhere
  - Public spaces, malls, offices, retail stores, public transport, ….
  - More cameras than eyes to monitor video

- Long-standing goal of *computer vision*
  - Let computers analyze all the video
  - Alert the user only when anything of interest is detected

} *Video Analytics (VA)*

Cost-effective platforms for Video Analytics



1) Server-based: Adds VA capability to existing video infrastructure

2) Edge-based: Smart network cameras transmit only events of interest and the associated meta-data

Challenges

- Application performance
  - Real-time detection of events
  - Low tolerance for false alarms or misses
- Algorithm
  - Unknown camera viewpoint
  - Unconstrained, varying scene illumination
  - Environmental factors (rain, camera movement, …)
  - Minimal user input
- Compute platform
  - Low power, low cost, high performance
  - Programmable for diverse vision functions
  - Many peripherals to interface other devices
  - High resolution video at 30fps

Example: Moving Object Segmentation

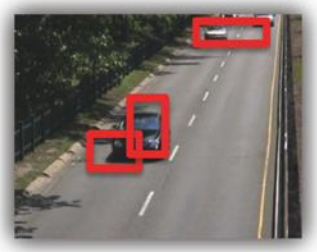
Common question in video management: is anything happening?

The Role of Moving Object Segmentation

SENSE

OBSERVE

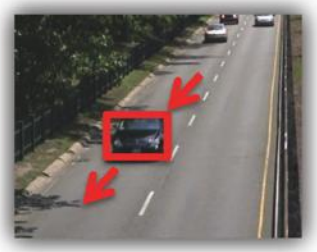
Moving Object Segmentation



DECIDE

Classifier

Tracker



RESPOND

Rules Engine
(Domain-based)

→ CAR COUNT

→ SPEEDING

An Approach for Moving Object Segmentation

VIDEO

Background Subtraction

foreground mask

Morphological Operations

smooth foreground mask

Connected Components

OBJECTS

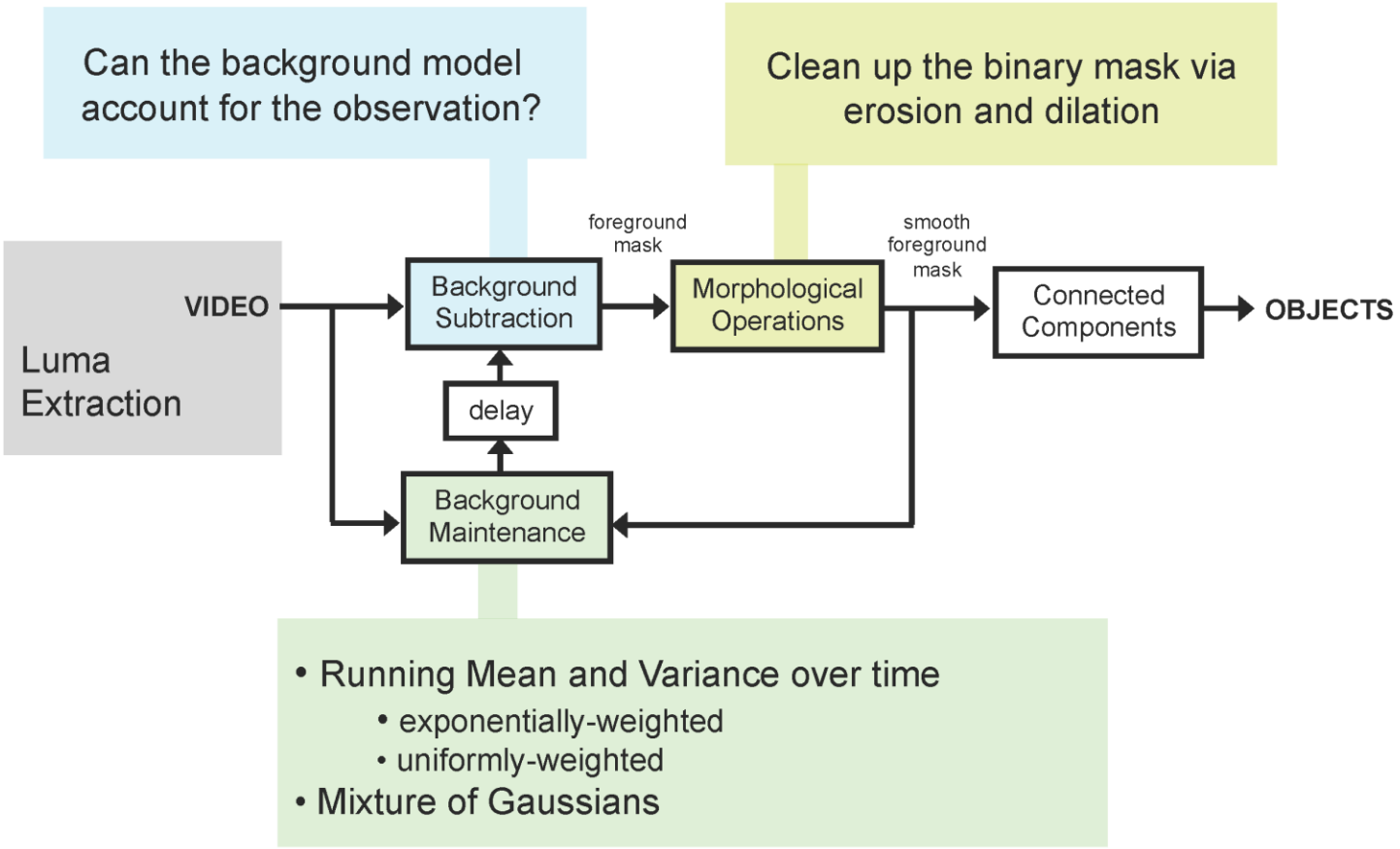
delay

Background Maintenance

BACKGROUND MODEL

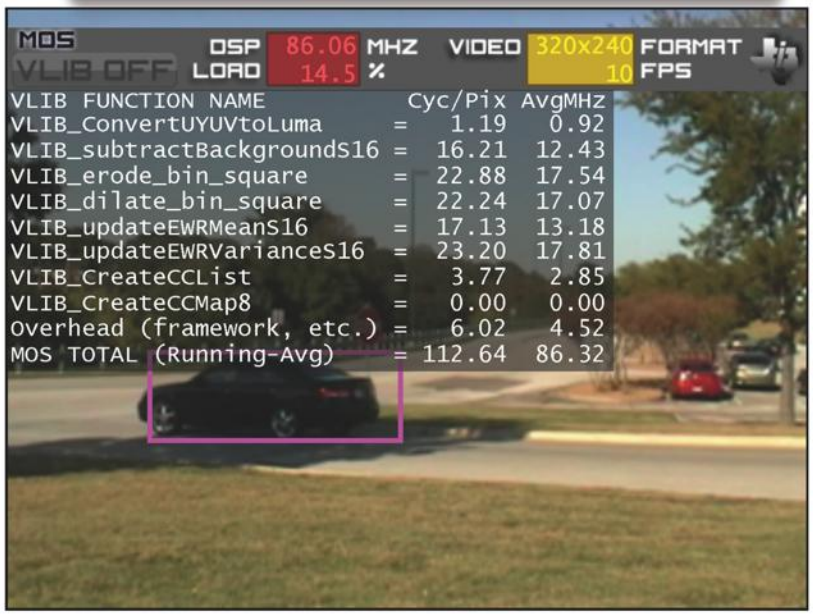
Common bottleneck for requirements including video analytics and intelligent occupancy sensing

Algorithms for Moving Object Segmentation

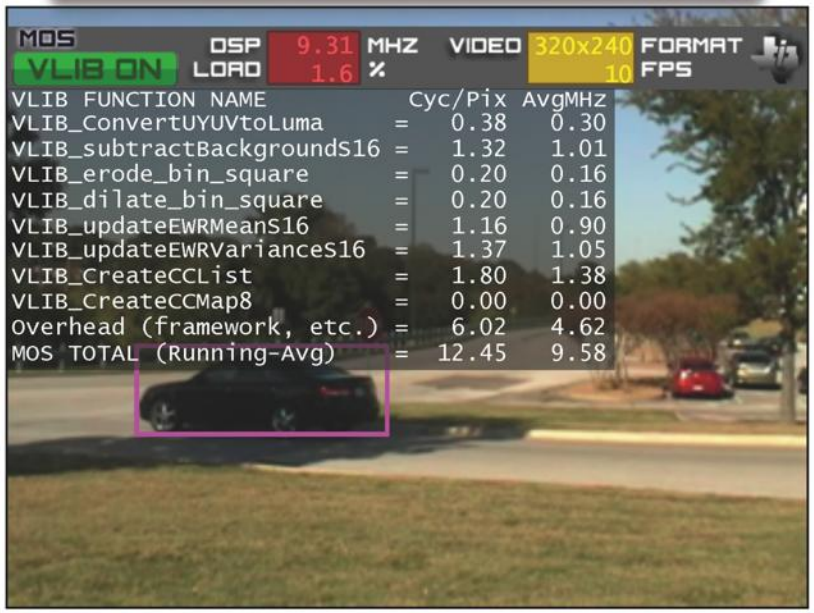


Need for efficient embedded implementation

VLIB off: utilizing 14.5% total processing power, 86 MHz



VLIB on: utilizing 1.6% total processing power, 9 MHz



DSP-optimized Vision Library (VLIB) provides 10x performance improvement over C code

Moving Object Segmentation performance on a TI DSP

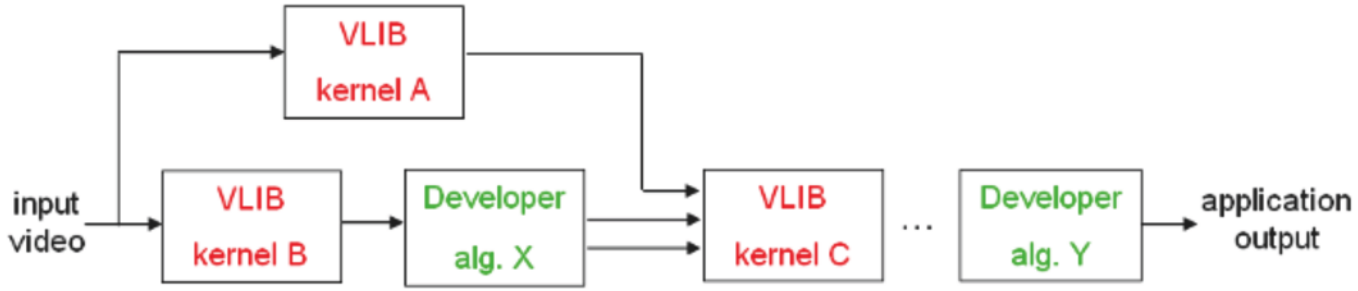
C64x+ MHz requirements (*)

	@ 5 fps	*@ 10 fps*	*@ 30 fps*
QVGA	**5**	**9**	**27**
CIF	**6**	**11**	**36**
VGA	**18**	**36**	**108**
D1	**21**	**41**	**123**
HD 720p	**54**	**110**	**333**
HD 1080p	**122**	**245**	**735**

(*) 16-bit precision, single-Gaussian model @ *QVGA resolution, VLIB turned ON, 50% overhead assumed*

Embedded algorithm design considerations

1. **High granularity**: small & well-understood operations.



2. **Fixed-point**: optimal use of Single-Instruction-Multiple-Data features of embedded DSPs.

```
int  VLIB_recursiveFilterHoriz1stOrder(
   char        *out,
   char        *in,
   int         width,
   int         height,
   short       weight,
   char        *boundaryLeft,
   char        *boundaryRight,
   char        *buffer);
```

3. **Data traffic**: Direct Memory Access friendly APIs.

VLIB functions by category

- **Background modeling & subtraction**
  - Luminance Extraction from YUV:422
  - Exponentially-Weighted Running Mean & Var.
  - Uniformly-Weighted Running Mean & Var.
  - Statistical Background Subtraction
  - Mixture of Gaussians Background Modeling & Subtraction
  - Morphological Operations (Erosion & Dilation)
  - Connected Components Labeling
- **Feature extraction**
  - Harris Corner Score (7x7)
  - Hough Transform for Lines
  - Histogram Computation for Integer Scalars
  - Histogram Computation for Multi-Dim. Vectors
  - Weighted Histogram for Integer Scalars
  - Weighted Histogram for Multi-Dim. Vectors
  - Legendre Moments
  - Canny Edge Detection
    - Smoothing
    - Gradient computation
    - Non-maximum suppression
    - Hysteresis
- **Low-level pixel processing**
  - Color Conversion YUV:422 interleaved to
    - YUV planar
    - RGB
    - LAB
    - HSI
  - Integral image
  - Image Pyramid (2x2 block averaging)
  - Non-Maximum Suppression (3x3, 5x5, and 7x7)
  - Gradient Image Pyramid (5-tap)
  - Gaussian Image Pyramid (5-tap)
  - First-Order Recursive IIR filters (horiz. & vertical)
  - SAD-based disparity for stereo
- **Tracking, recognition, etc.**
  - Lucas-Kanade Feature Tracking (7x7)
  - Kalman Filtering
  - Nelder-Mead Simplex optimization
  - Bhattacharya distance

Challenges (revisited)

- Application performance
  - Real-time detection of events
  - Low tolerance for false alarms or misses
- Algorithm
  - Unknown camera viewpoint
  - Unconstrained, varying scene illumination
  - Environmental factors (rain, camera movement, …)
  - Minimal user input
- Compute platform
  - Low power, low cost, high performance
  - Programmable for diverse vision functions
  - Many peripherals to interface other devices
  - High resolution video at 30fps

Demonstrated "Moving Object Segmentation" and its implementation through VLIB functions on an embedded SoC

Thank you!

G. Dedeoglu, B. Kisacanin, D. Moore, V. Sharma, and A. Miller, "**An optimized vision library approach for embedded systems**", *Proceedings of the Seventh IEEE Workshop on Embedded Computer Vision, pp. 8-13, 2011 (CVPR).*
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5981731

http://www.ti.com/vlibrequest