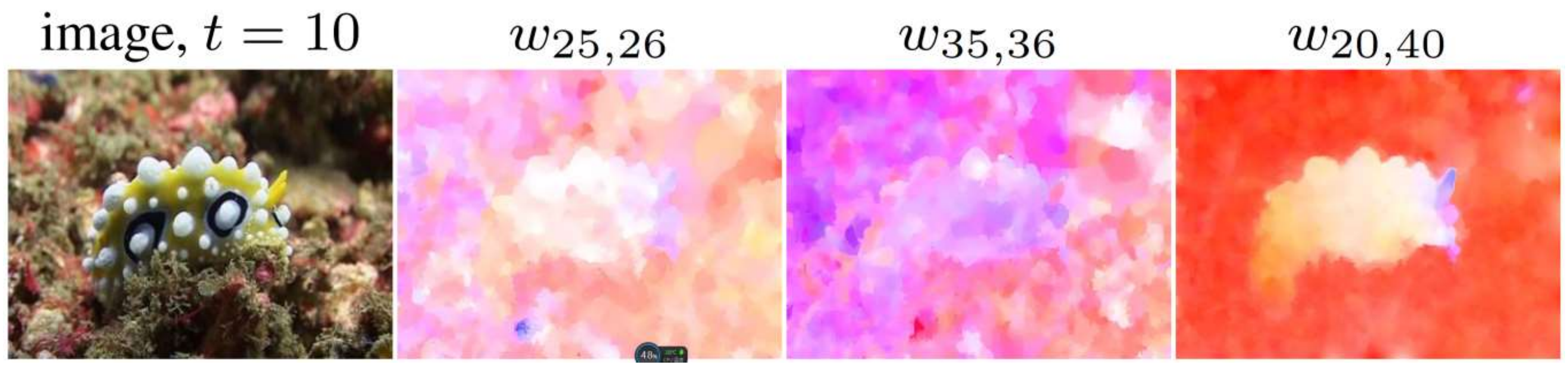
Minimum Delay Moving Object Detection

Dong Lao, Ganesh Sundaramoorthi†
† King Abdullah University of Science and Technology

ABSTRACT

We present a general framework and method for detection of an object in a video based on apparent motion. The object moves relative to background motion at some unknown time in the video, and the goal is to detect and segment the object as soon it moves in an online manner. Due to unreliability of motion between frames, more than two frames are needed to reliably detect the object. Our method is designed to detect the object(s) with minimum delay, i.e., frames after the object moves, constraining the false alarms. Experiments on a new extensive dataset for moving object detection show that our method achieves less delay for all false alarm constraints than existing state-of-the-art.



Composition of warps across frames produces a strong motion sig-nal. [Left]: Image. [Middle two]: Optical flows between adjacent frames at two instances show that the object is not clearly visible. [Right]: Composition of warps between frames 10 and 40 shows the object is clearly visible. How many frames does it take to reliably detect the moving object? Our method addresses this question.

MULTIFRAME MOTION SEGMENTATION

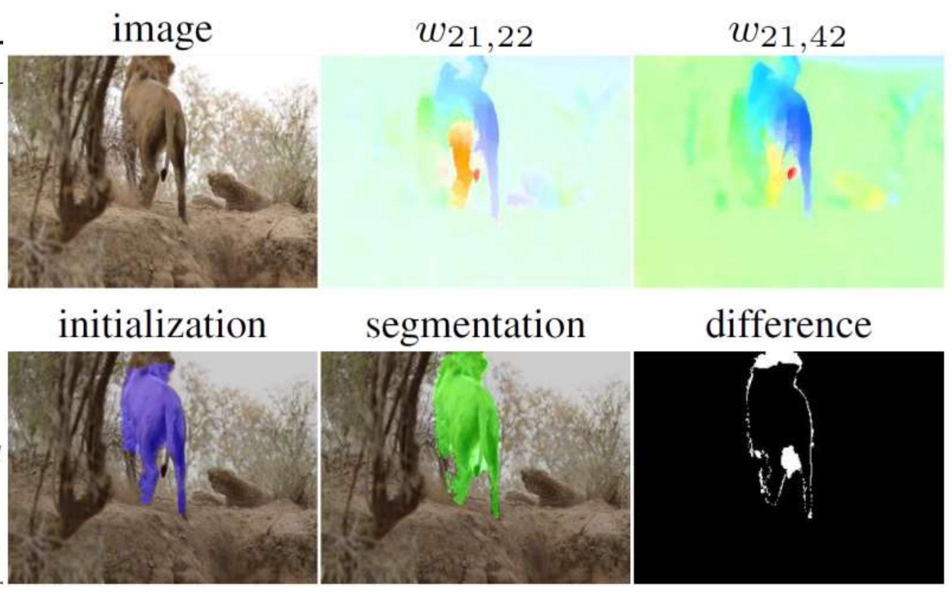
We minimize the following energy for multiframe motion segmentation:

$$E_{seg}(\{R_s^i\}_{i=0}^{n-1}) = \underbrace{-\sum_{i=0}^{n-1} \log[p(R_s^i)]}_{\text{Region shape prior}} + \sum_{i=0}^{n-1}\int_{R_s^i} \underbrace{[1-m(x)]f^i(x)}_{\text{Motion term}} \underbrace{- m(x)\log p_{R_s^i}(I_s(x))}_{\text{Color histogram term}}\,\mathrm{d}x$$

$$\text{where } f^i(x) = \sum_{t=t_1}^{t_2} \rho_i[I_{t+1}(w_{s,t+1}^i(x)) - I_t(w_{s,t}^i(x))] \det \nabla w_{s,t}(x)$$

Algorithm 1 *Multiframe motion segmentation*

```
Input: I_{t1:t2} and s ∈ [t1, ..., t2]
// initialize R^i_s for gradient descent of E_seg
Compute Classic-NL warp w^{NL}_{t,t+1} : Ω → Ω, ∀t
Compute w^{NL}_{s,t1}, w^{NL}_{s,t2} by composing warps
Use w^{NL}_{s,t1}, w^{NL}_{s,t2} as channels to segment
repeat // gradient descent of E_seg for R^i_s
    Propagate R^i_s frame-wise to form R^i_t, ∀t
    Solve for Sobolev warp w^i_{t,t+1}
    Compute w^i_{s,t} for t ∈ {t1, ..., t2}\{s}
    Compute f^i and update R^i_s by gradient step of E_seg
until R^i_s does not change between iterations
Propagate R^i_s to form R^i_{t2}
return {R^i_{t2}}^{n-1}_{i=0} as the segmentation in frame t2
```



Demonstration of multiframe motion segmentation. [Top row]: an image in the sequence, optical flows between adjacent frames, and composed optical flow. [Bottom]: Initialization to motion segmentation, final segmentation, and the difference between the two.

QUICKEST MOTION DETECTION

The goal of *Quickest Detection* is minimizing the delay with guarantees on reliability of the detections. We combine the QD statistics model with our motion model as follows:

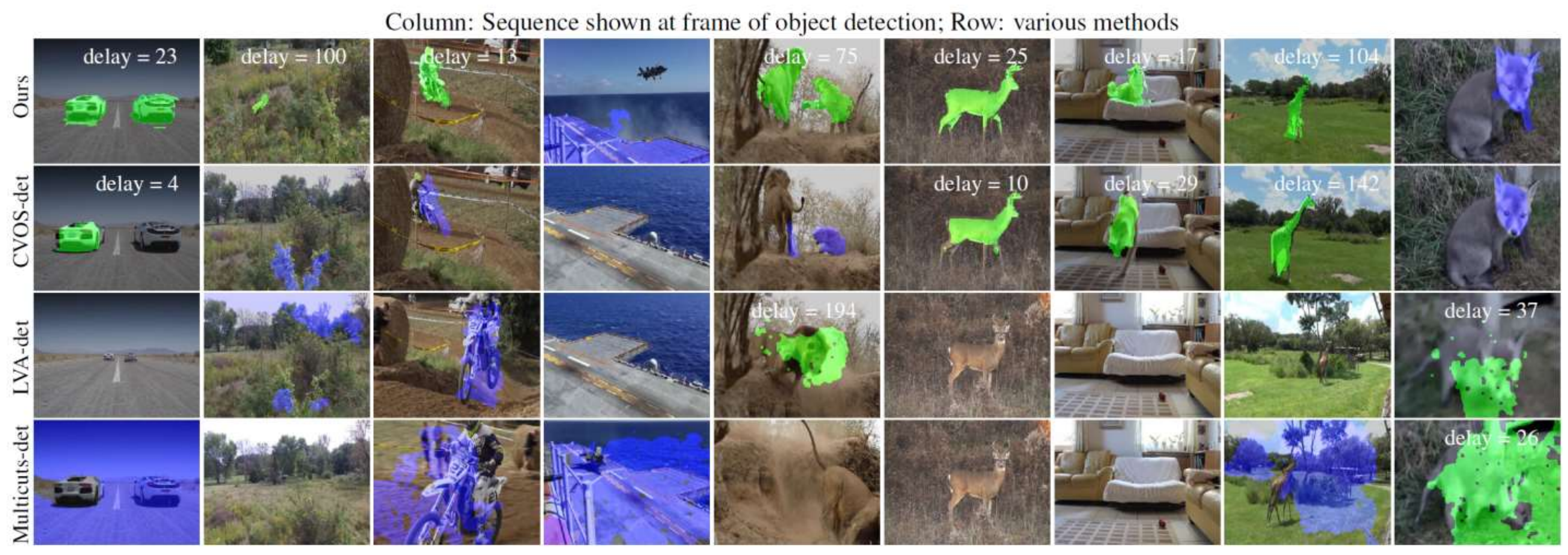
$$p_0(\mathbf{I}_{i:i+1}|v_i^0) \propto \exp\left\{-\int_\Omega \rho_0[I_{t+1}(w_{t,t+1}^0(x)) - I_t(x)]\,\mathrm{d}x\right\}$$

$$p_1(\mathbf{I}_{t:t+1}|\{v_t^i, R_t^i : 0 \le i < n\}) \propto \exp\left\{-\sum_{i=0}^{n-1}\int_{R_t^i} \rho_i[I_{t+1}(w_{t,t+1}^i(x)) - I_t(x)]\,\mathrm{d}x\right\}$$

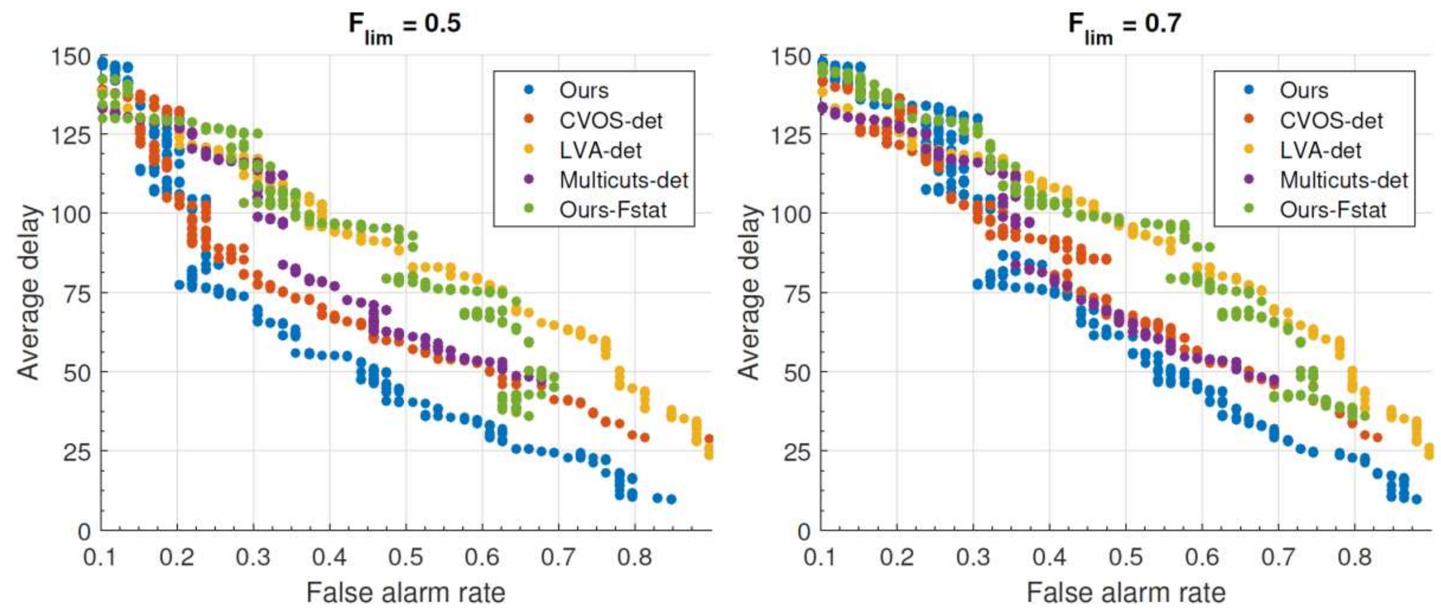
We threshold $\Lambda_t = \max_{1 \le t_c < t} \max_\theta \prod_{s=t_c}^{t} \frac{p_\theta(X_s)}{p_0(X_s)}$ to obtain a detection time and corresponding segmentation.

RESULTS

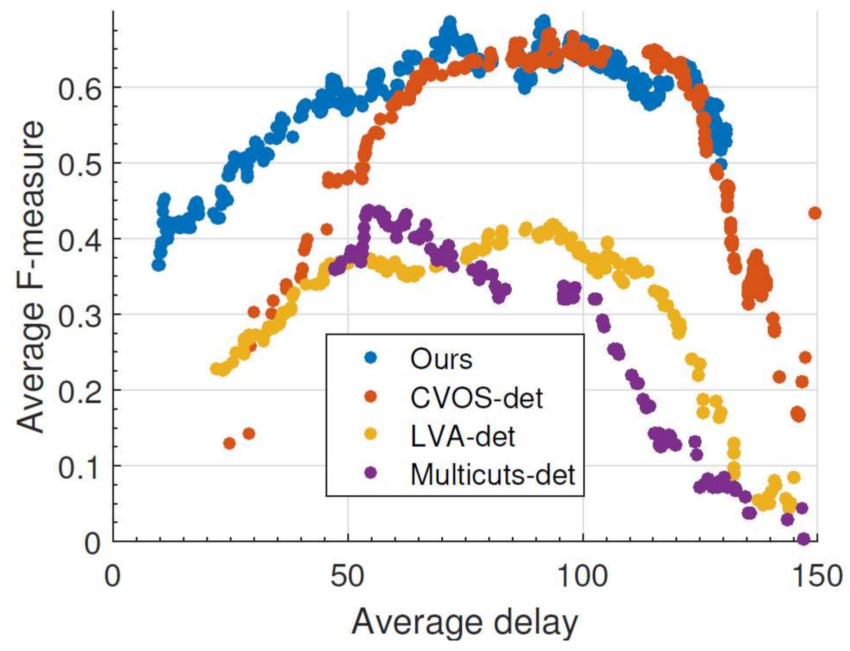
Since there are no datasets that are explicitly designed for detection of moving objects, we created our own of 78 videos, varying from 100 to 800 frames. The camera moves, and the object moves (differently than the background) at some unknown frame. The videos may consist of a single or multiple objects.



Representative detections for methods tested. We show visualizations of the detections (or non-detections) for each of the competing methods, each operating at a false alarm rate of 0.3. Green masks indicate a correct detection, while purple masks indicate a false alarm. The delay at the detection is indicated, if the detection is correct. No segmentation result indicates the method did not detect (maximum delay). Results illustrate our method achieves less delay on average.



Delay versus false alarm curves. All moving object detectors are compared. [Left]: Threshold for measuring false alarms is 0:75 and [Right]: 0:5. Results show ours method has the least delay.



Segmentation accuracy at detection time. All moving object detectors are compared in terms of their average F-measure to ground truth at detection time. Results show our method has highest accuracy at all levels of delay.

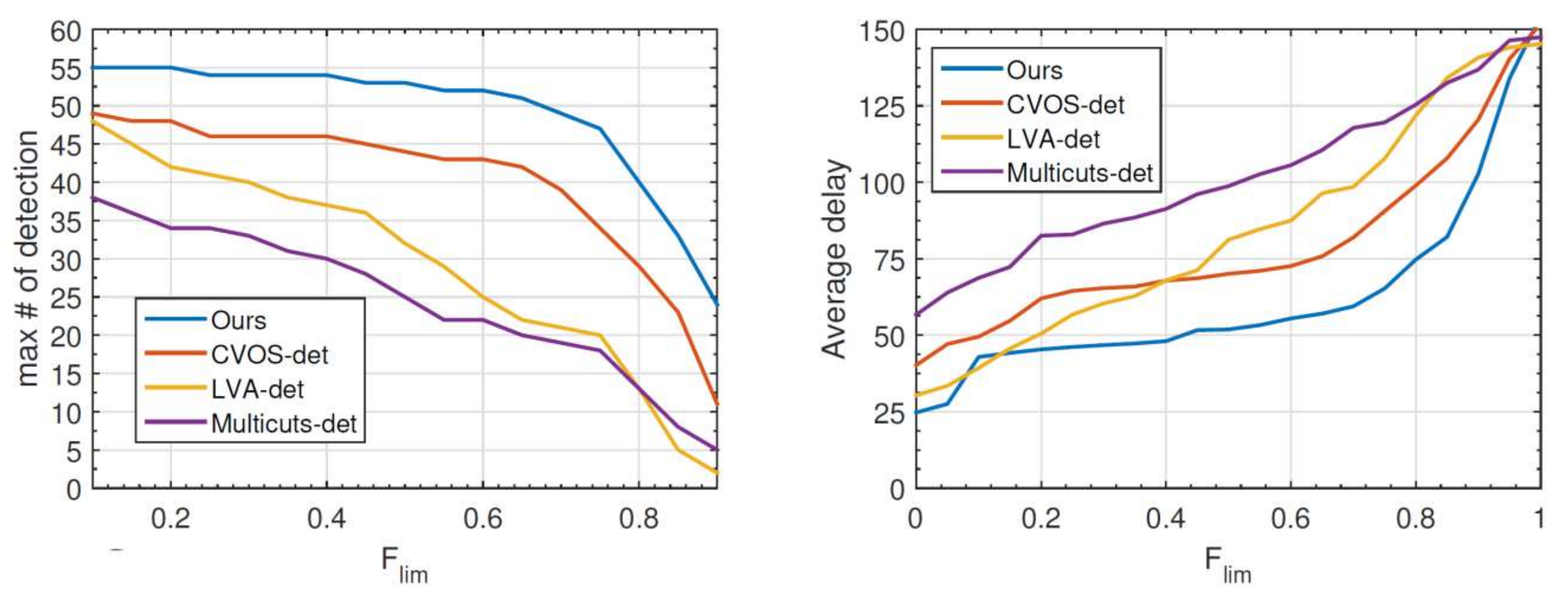


Figure 8: Correct detections and delay of ideal detection mechanisms. [Left]: Correct detections versus the threshold for measuring false alarms. Under any measure of false alarms and ideal detection mechanisms, our method achieves more detections. [Right]: We also achieve less delay.