Sandia National Laboratories

SAND2014-1489C

LDRD LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Real-Time Change Detection for Wide Area Surveillance

Conference on Data Analysis (CoDA)
06-Mar-2014   Santa Fe, NM

Katherine Simonson and Tian Ma
Sandia National Laboratories

Contact Author: Katherine Simonson
email: kmsimon@sandia.gov
tel: 505-844-6551

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04094AL85000.

Outline

- Background and Mission Constraints
- Mathematics of the Detector
- A Few Examples
- Summary and Conclusions

Research Goal

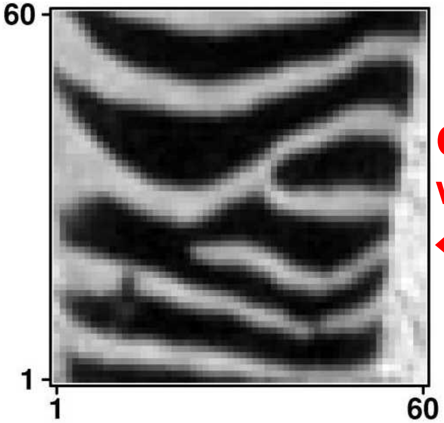
- **Our goal is the *automatic detection of small changes* in wide area surveillance.**
  - **We work primarily with *low-resolution, staring, radiometric sensors*, which are subject to significant jitter.**
  - **Frame rates up to 75 Hz; *algorithms must run causally in real time*.**
- **We *work directly with real-world data* from deployed, operational, surveillance systems.**
  - **As well as video sequences from a range of unclassified sources.**
- **We are interested in *detecting all physical change in the scene*, no matter how small.**
  - **While rejecting variation due to sensor-related artifact, including pointing drift, jitter, noise, pixel irregularities, and specularities.**

Background Subtraction

- The standard approach to *change detection involves some form of subtraction*:
  - To detect new energy at time *t*, subtract from the frame taken at *t* an estimate of the "background" energy in the scene prior to this time.
  - The background estimate may be a single prior frame or a more complex function evaluated over a window of recent frames.
- If the current frame is not properly registered to the background, large values in the difference frame may be caused by intensity gradients in the scene, rather than true (physical) change.
- It follows that *change detection in a high jitter environment is particularly challenging!*

Mis-Registration

Frame 1

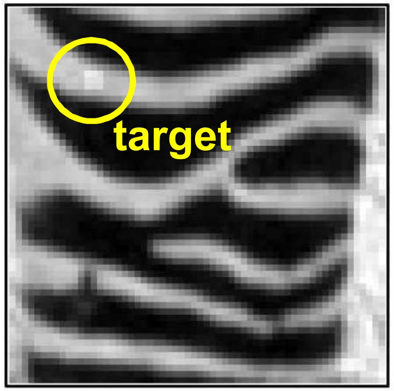


Crop and add white noise

Full Image

Crop, shift (r,c)+0.5, add white noise and 3x3 target

Frame 2

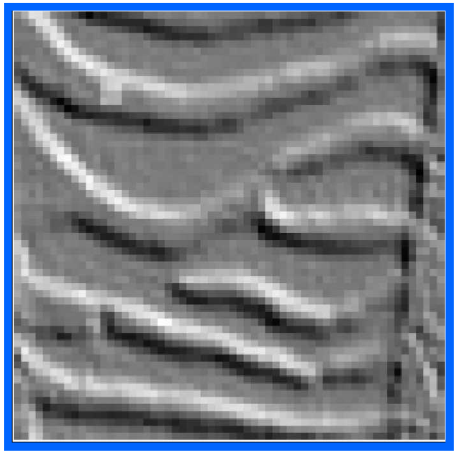


The difference between frames slightly out of alignment is dominated by scene gradients larger than the target change.

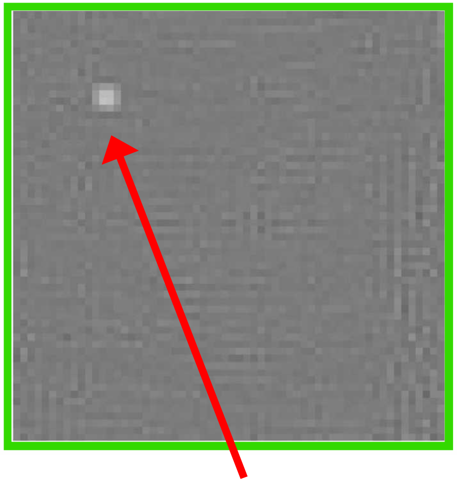
Interpolating the second frame into alignment with the first blurs the target signal.

The two difference frames are plotted in the same greyscale.

Frame 2 – Frame 1, Unregistered



Frame 2 – Frame 1, Registered



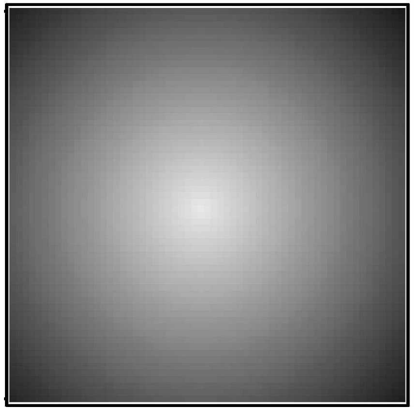
target signal, blurred

Sensor Artifact

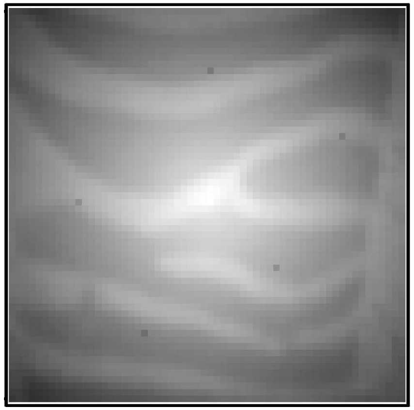
Artifacts in pixel space challenge solutions based on scene registration!

A bias surface was added to the original frames, and reduced responsiveness was simulated in 5 pixels.

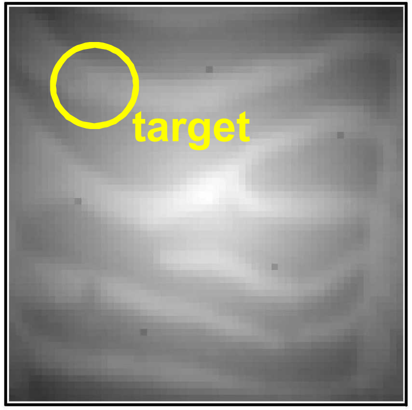
Bias Surface



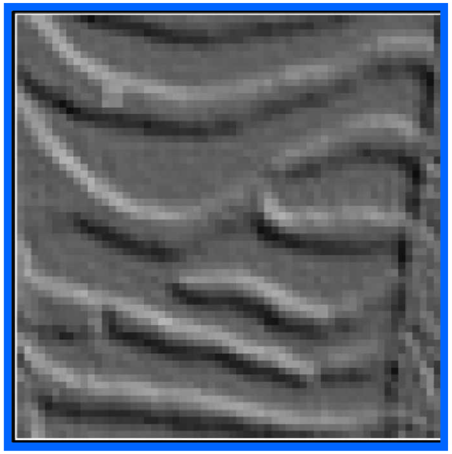
Frame 1 + Bias, Reduced Response



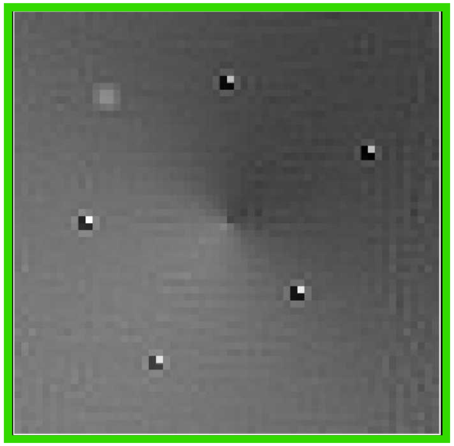
Frame 2 + Bias, Reduced Response



Difference Frame, Unregistered



Difference Frame, Registered



When Frame 2 is translated to register with the scene of Frame 1, the defects move out of alignment, creating large apparent changes in the difference frame.

All such defects must be known and corrected for prior to scene registration.

Algorithm Approach

- *Frame registration cannot solve the jitter problem* in real time:
    - Registration to a small fraction of a pixel is required, but this precision is often not achievable at high frame rates for low-quality data.
    - Even if jitter-induced offsets are known perfectly, all sensor artifacts (fixed pattern noise, self-emission, over- or under-responsive pixels) have to be corrected prior to frame transformation. This may not be feasible for gradually varying artifacts.

- Our approach does not require registration, instead relying on two separate statistical models for variations in pixel intensity.
    - *The temporal model* handles pixels that are naturally variable due to sensor noise or moving scene elements, along with jitter displacements comparable to those observed in the recent past.
    - *The spatial model* captures jitter-induced changes that may or may not have been observed previously.

Normalized Differences

- For each pixel (*k*, *h*) at time *t*, ***we test whether the observed intensity is consistent with the spatial and temporal models***. The decision is based on simple normalized differences.

$$Z(k,h;t) = \frac{X(k,h;t) - B(k,h;t)}{S(k,h;t-1)}$$

- ***X*** = pixel (***k, h***)'s intensity at time ***t***,
- ***B*** = current background estimate,
- ***S*** = current standard deviation estimate

- A large (absolute) value of ***Z(k, h; t)*** means the observed pixel intensity is outside the range anticipated under the current model.

Decision Logic

$$Z(k,h;t) = \frac{X(k,h;t) - B(k,h;t)}{S(k,h;t-1)}$$

- **Normalized differences, $Z_{SPATIAL}$ and $Z_{TEMPORAL}$ are computed using the same background *B*, but *two different standard deviation estimates*.**

- **If min { | $Z_{SPATIAL}$|, | $Z_{TEMPORAL}$| } exceeds a fixed threshold, the observed value of the pixel at time *t* is inconsistent with both models, and a candidate detection occurs.**

  - A one-sided test may be applied if, e.g., only positive deviations are of interest.

- ***Depending on the characteristics of the target changes sought*, downstream logic may be employed to reduce the false alarm rate:**

  - *Area filtering:* Require detection in at least *N* connected pixels.

  - *Duration filtering:* Require detection in at least *M* consecutive frames.

Background Estimation

- **Scene background is estimated using *adaptive subspace projection.***
  - The goal is to capture the covariance structure of a sequence of frames in a low-dimensional, orthogonal subspace.
  - The general approach applied to the jitter problem dates to 1983 [1]. Many algorithms exist; some (e.g., FAPI [2]) can update in real time.
  - ***You do not need to estimate, store, or decompose a data covariance matrix*** **!** Storage of past frames is not required.

- **New frames are projected into the jitter subspace; pixelwise change is measured via the projection residuals.**
  - To track gradual change in a scene (e.g., cloud motion), the subspace is updated after each frame. The decay rate, $\beta \in [0,1]$, is tunable.

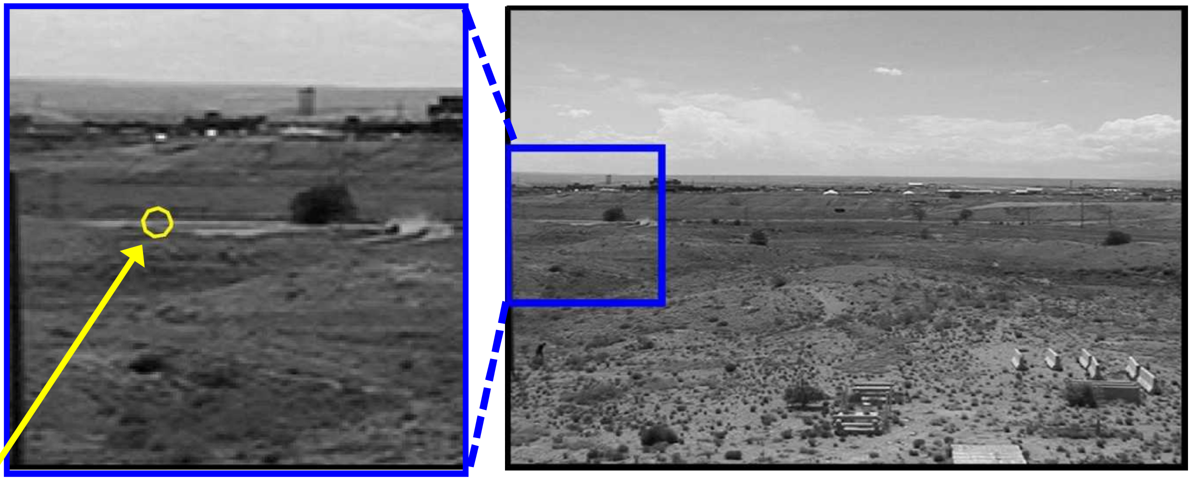
- **When a strong detection occurs, the background update rate can be slowed down for the affected pixels.**
  - Prevents target energy from being immediately absorbed into the background.

1. P.E. Barry and M. Klop, "Jitter Suppression: A Data Processing Approach," *Proc. SPIE*, vol 366, pp. 2-9, 1983.
2. R. Badeau, B. David, and G. Richard, "Fast Approximated Power Iteration Subspace Tracking," *IEEE Trans. Signal Proc.*, vol. 53, no. 8, pp. 2931-2941, 2005.

Background Estimation: Example

The highlighted pixel lies along a road, and is subject to change due to both camera jitter and passing traffic.

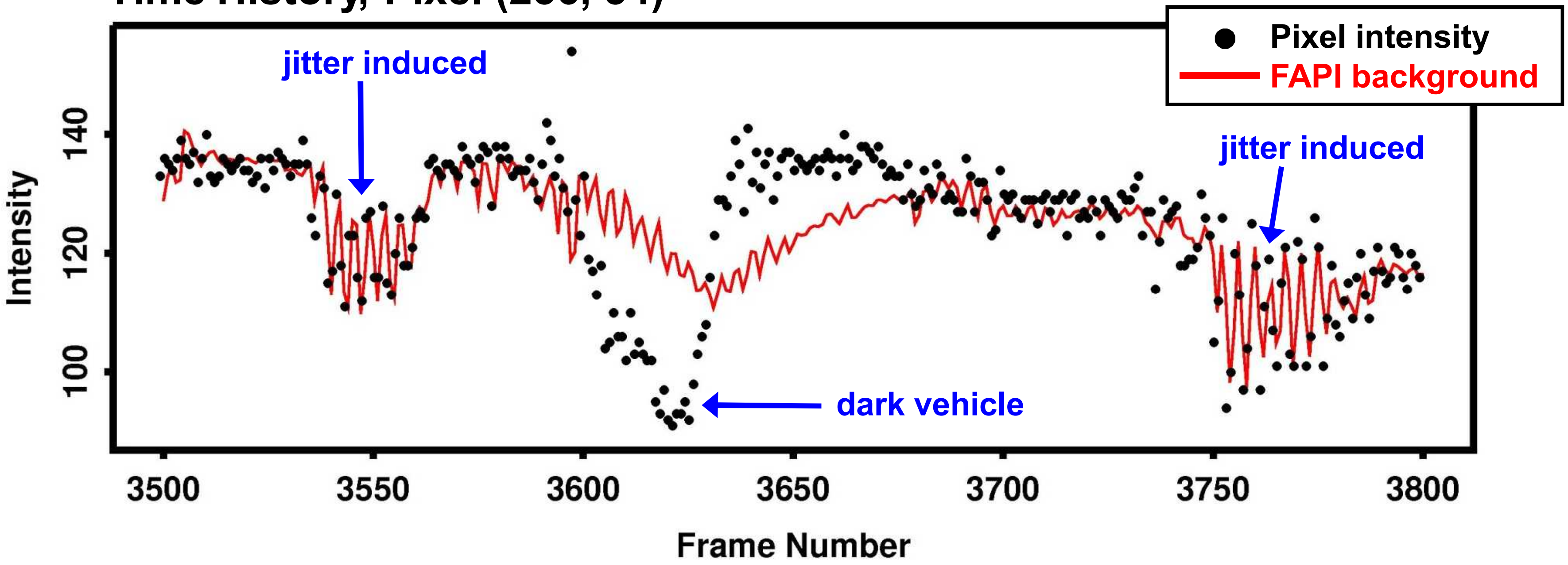
The FAPI background estimate tracks jitter closely, but gives large residuals when a dark vehicle moves through.



Expanded View

Full Scene

Time History, Pixel (256, 54)



Temporal Variances

- **"Temporal" estimates of pixel variance are based on a recent time window of projection residuals. They are computed as follows:**

  1. Initialize with the sample variance over the first *n* frames, *V*(*k*,*h*; *n*).

  2. For subsequent frames, update using:

$$V(k,h;t) = (1-\gamma)\left[X(k,h;t) - B(k,h;t)\right]^2 + \gamma\, V(k,h;t-1)$$

- **Forgetting factor $\gamma \in [0,1]$ determines how rapidly the filter responds to new energy.**

  - As with the background estimate, the temporal variance estimate for any pixel showing a strong detection can be updated more slowly.

Spatial Estimation: Motivation

- As long as the jitter distribution is relatively stable, the temporal approach to variance estimation provides reasonable scale factors.

- For non-stationary jitter, temporal estimates are inadequate: when jitter increases, false alarms occur along scene gradients.

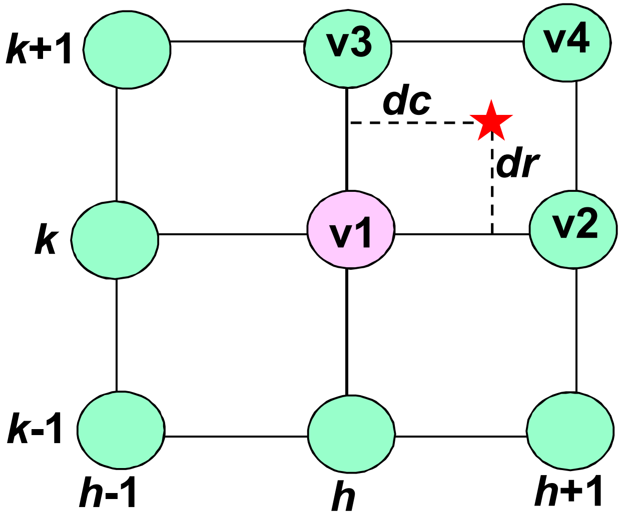
  - Subspace projection alone does not solve this problem !

Key Observation: You do not need to observe line-of-sight jitter to predict which pixels will be influenced !

- We have developed a new mathematical concept for pixel variance estimation. Our "spatial" approach produces estimates that are robust to non-stationary jitter, based on a single frame.

Bilinear Interpolation

- The method operates over a grid of conditional expectations in the vicinity of each pixel.



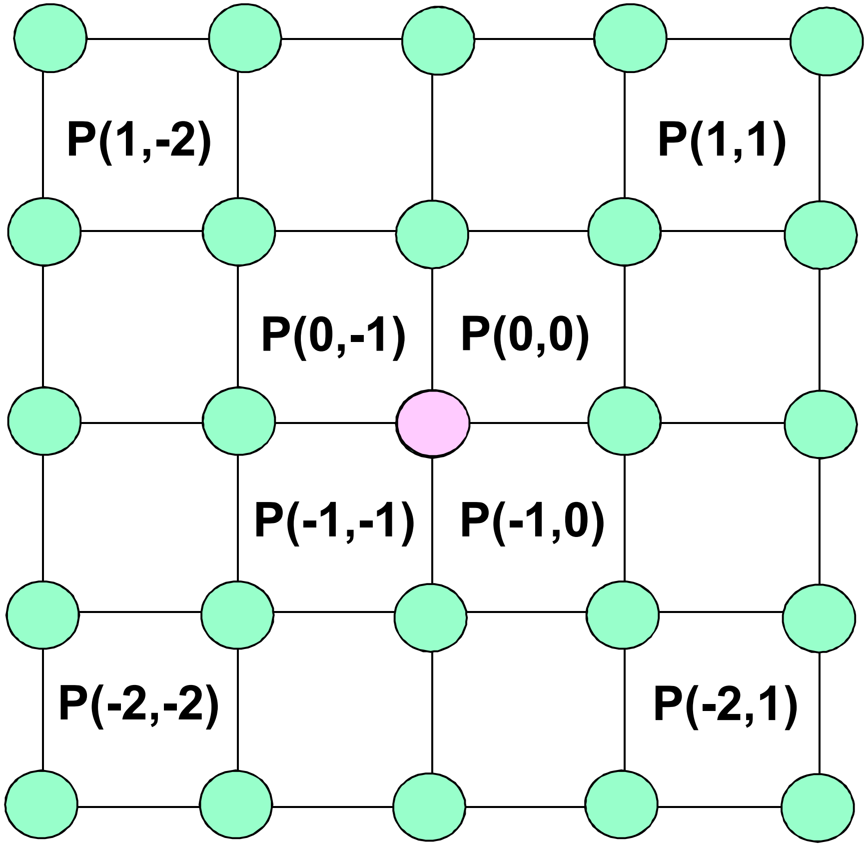
- At time *t*-1, define:
  v1 = value at pixel (*k*,*h*)
  v2, v3, v4 = values at nearby pixels

- *If we knew that jitter between times t-1 and t was exactly dr rows and dc columns*, we could use bilinear interpolation to estimate the background at pixel (*k*,*h*) at time *t*:

$$E(k,h;\ t) = v1 + dr\,(v3-v1) + dc\,(v2-v1) + dr \cdot dc\,(v1+v4-v2-v3)$$

- If (*dr*, *dc*) is unknown, we can *use its statistical distribution to estimate the mean and variance* of each pixel at time *t* as a function of pixel values at time *t*-1 (or other previous frame).

Conditional Expectation



- For each “cell” near (*k,h*), we use an ***assumed jitter distribution*** to compute:

1) The probability of jittering into this cell at time *t*, and:

2) The expected pixel value (and its square) at *t*, given jitter into this cell.

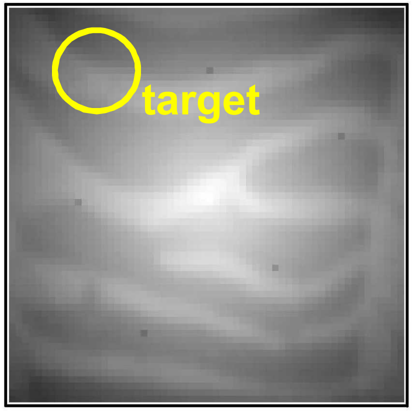
- After much algebra (see SAND report), we ***apply the Law of Total Probability to estimate the variance*** of each pixel at time *t*.

- Estimates computed in this manner are ***surprisingly robust to mis-specification of the jitter distribution***: They scale roughly linearly with the jitter standard deviation parameter (sigma).

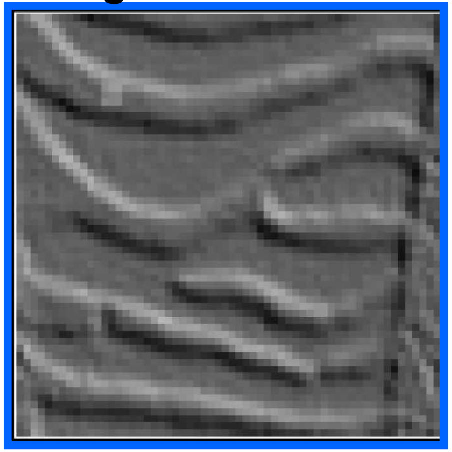
  - A good strategy is to set the sigma conservatively (based on the worst jitter expected) and re-scale on a per-frame basis.

Incorporating SSP and Spatial Variances

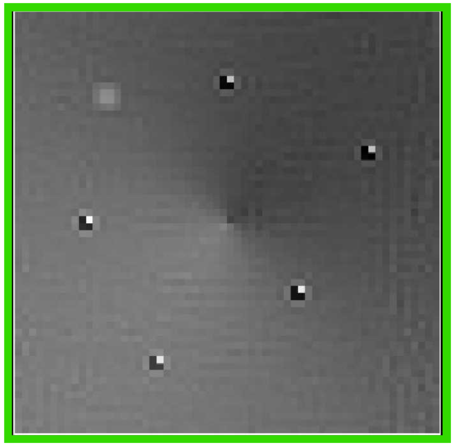
Frame 2 + Bias, Reduced Response



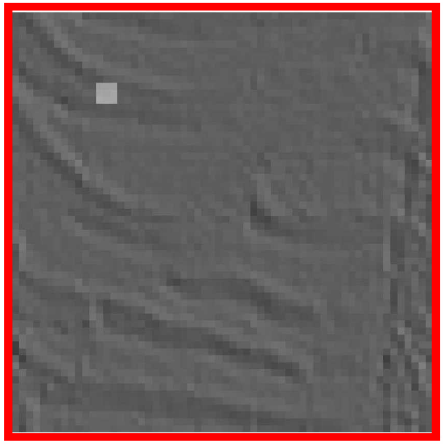
Difference Frame, Unregistered



Difference Frame, Registered



Raw SSP Residuals



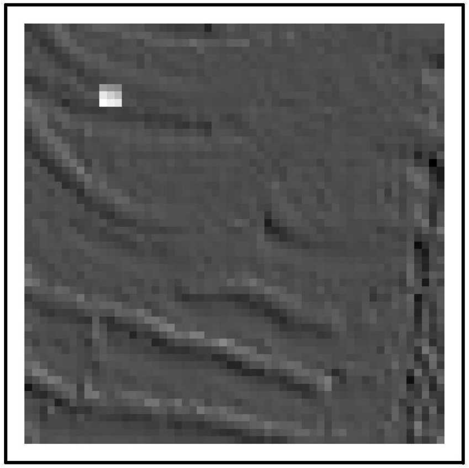
The principal subspace was estimated from 100 simulated jittered, noise-added versions of Frame 1 (with bias surface and reduced responsiveness).

SSP residuals show less scene structure than the unregistered frame differences, and exhibit no sensor artifact.

After division by spatial standard deviations, the nine target pixels have values between 1.51 and 4.55, larger than ALL non-target pixels.

Spatial StDevs

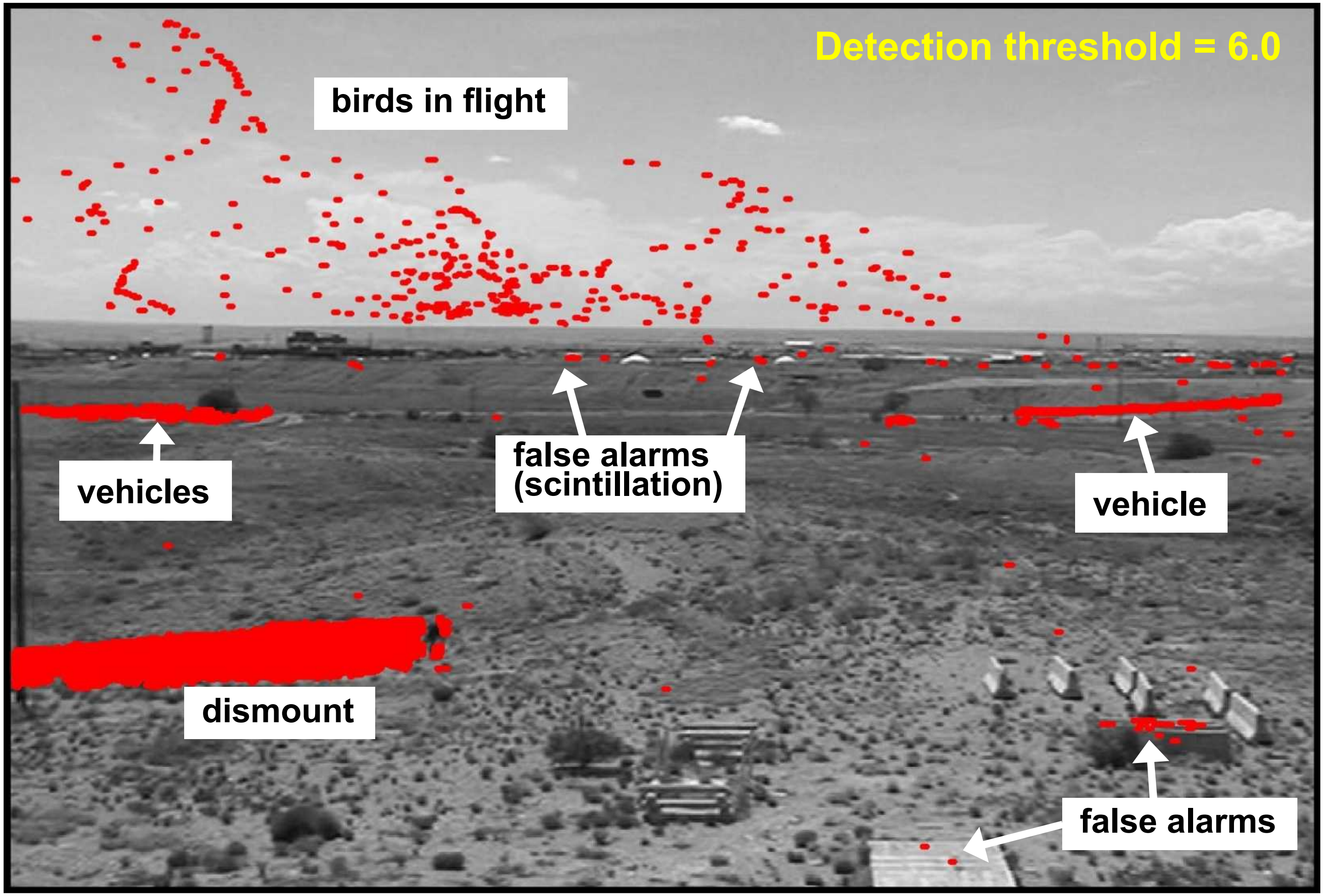
Normalized SSP Residuals



Example 1 – Kirtland AFB

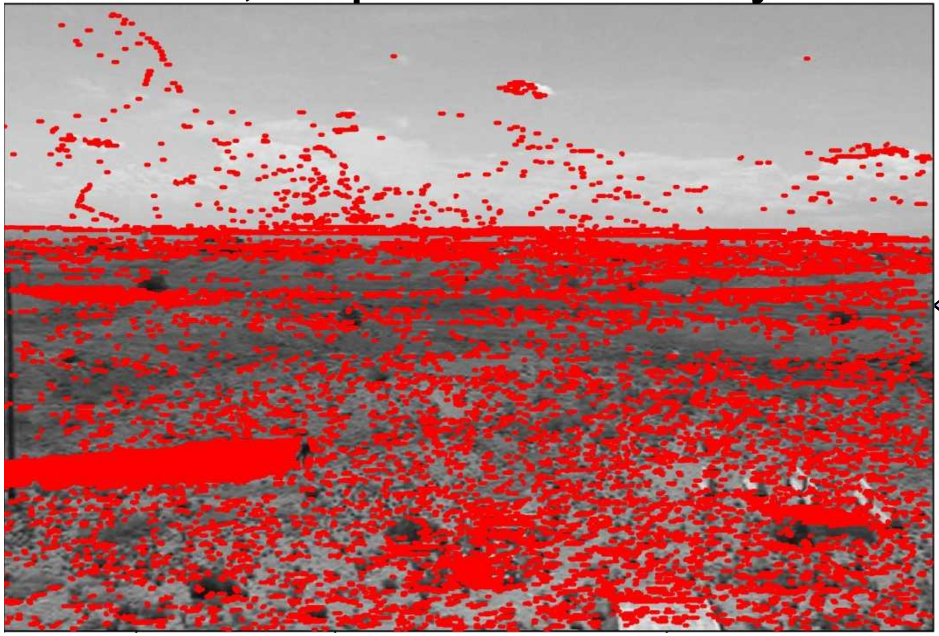
30 Hz video showing various activities near Sandia's robotic vehicle range.

Red dots show pixels with at least one detection in frames 2400 – 3800, using the dual-variance (spatial & temporal) model.



Example 1, Cont'd

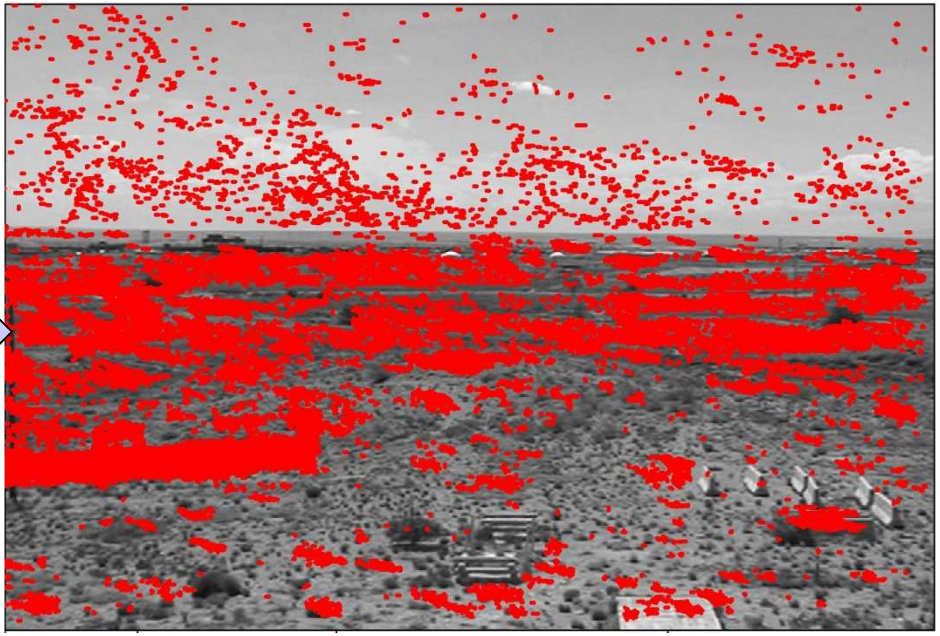
Detections, Temporal Variances Only



Frames 2400 – 3800
Detection Threshold = 6.0

When only temporal estimates of pixel variance are available, false alarms occur at scene edges: bright clouds, roads, vegetation, and the horizon.

Detections, Spatial Variances Only



When background differences are normalized with spatial standard deviation estimates only, sensor noise induces false alarms in relatively uniform parts of the scene.

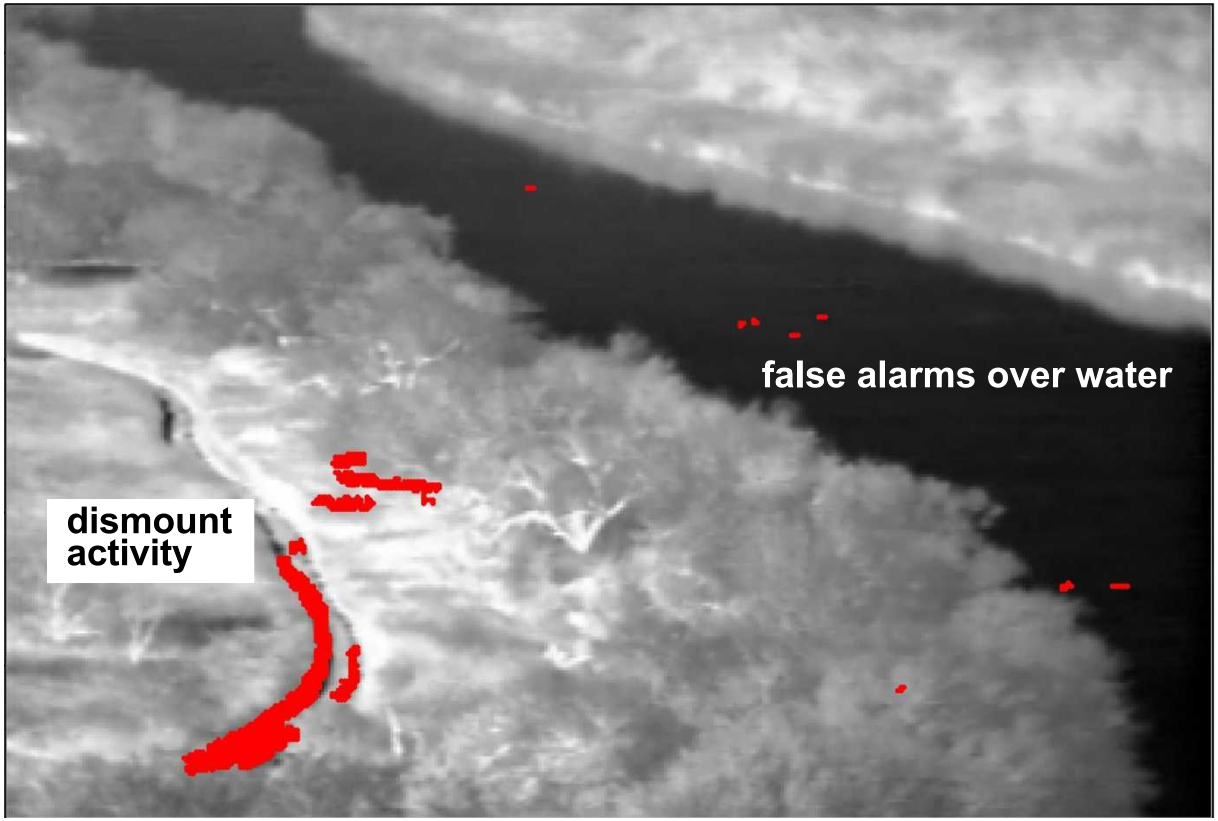
KAFB Video with Detections

Red boxes indicate pixel detections; no tracker is applied.

Example 2 – Border Camera Footage

10 Hz video sample from a pole-mounted infrared camera*; nighttime scene on the Texas/Mexico border.

Red dots show pixels with at least one detection in frames 500 – 1500, using the dual-variance model.



Two dismounts emerge from the vegetation along the river, return to the riverside, re-emerge, and proceed down the track and out of the scene.

At times, they are lost in the near-saturated pixels to the right of the track.

***Texas Border Sherriff's coalition and www.blueservo.net**

Border Video with Detections

Red boxes indicate pixel detections; no tracker is applied.

Run-Time Performance

- *Real-time implementations* of the detection algorithm described here have been utilized for a variety of applications over the past several years.

  - The software has been tested on frame sizes as large as 2k × 2k, at frame rates up to 75Hz.

- Scene background and temporal variance estimates are *efficiently updated after every frame.*

  - For large frames, adaptive subspace estimation (FAPI) processing runs on specialized GPU hardware.

- Spatial variance estimates are currently updated once per second.

  - Sufficient for slowly-changing background or gradual pointing drift.

  - We are planning upgrades to a higher refresh rate, to enable *robust change detection even in the presence of fast pointing drift*.

Summary

- **The algorithm outlined here is designed to provide robust change detection, even in the presence of platform jitter, pointing drift, and significant sensor artifacts.**

- **The three key elements are:**

  1. ***Background modeling*** **via adaptive subspace estimation;**
  2. ***Temporal variance estimates*** **to track historical change;**
  3. ***Spatial variance estimates*** **to model susceptibility to jitter and/or pointing drift.**

- **The approach has** ***proven performance in real-world operations*****.**

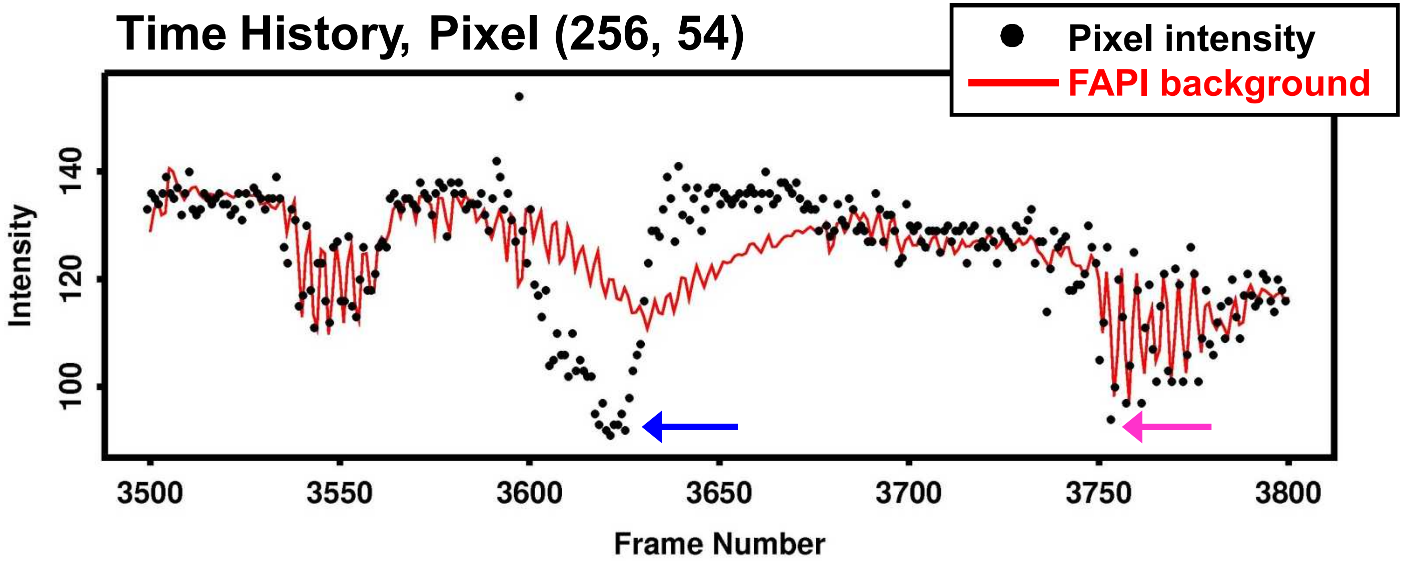
- **Sandia was granted a U.S. Patent for the spatial variance estimation technique.**

U.S. Patent No. 8,103,161, K.M. Simonson and T.J. Ma, "Estimating Pixel Variances in the Scenes of Staring Sensors," 24-Jan-2012.

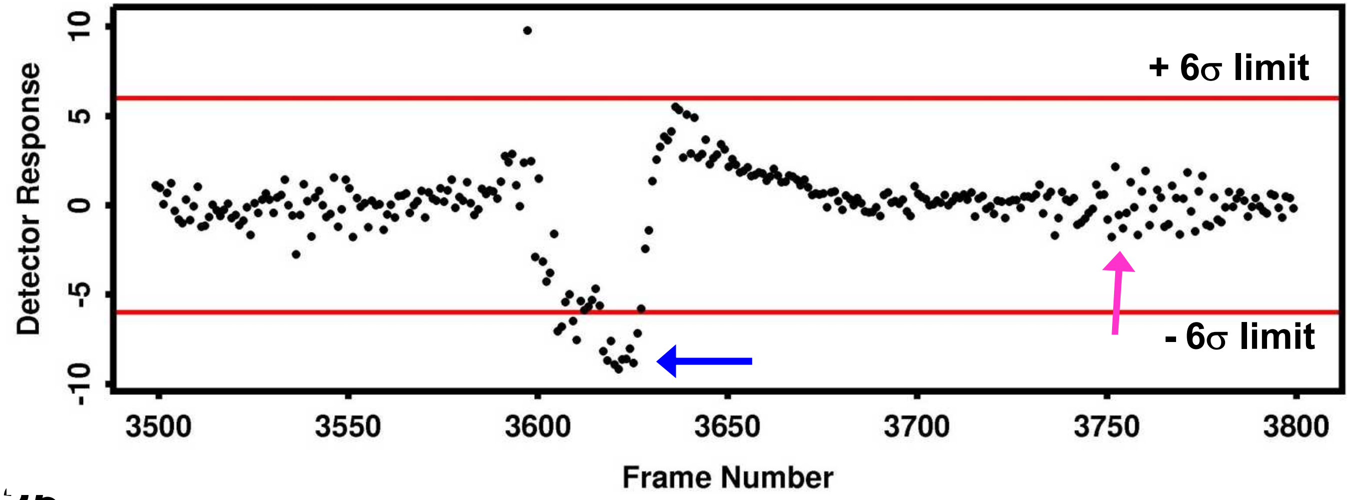
BACKUP SLIDES

Detector Response

Time History, Pixel (256, 54)



Detector Response, Pixel (256, 54)

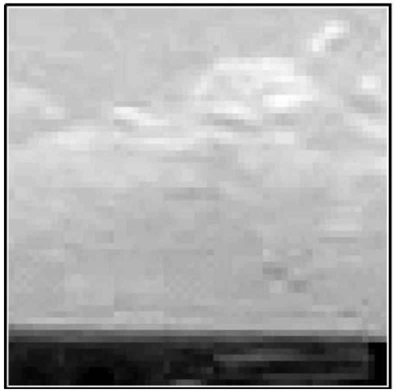


While the decreased intensity due to jitter (pink arrow) is almost as low as the drop due to a dark vehicle passing through the pixel (blue arrow), the detector responds differently to jitter and signal.

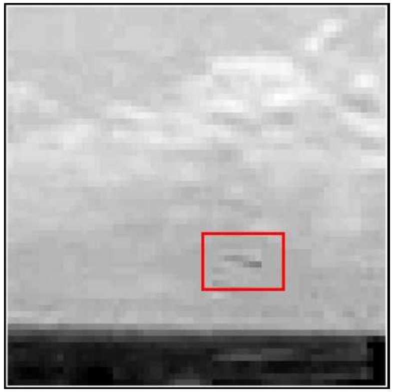
Birds, Birds, Birds!

Detected Pixels, frames 2891 - 2900

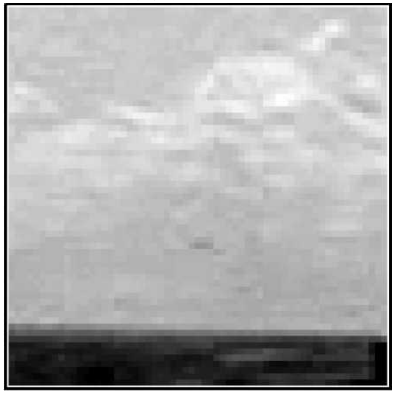
Frame 2891



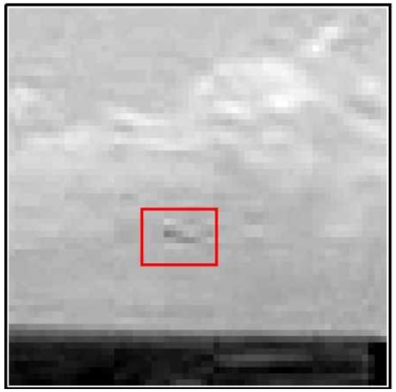
Frame 2892



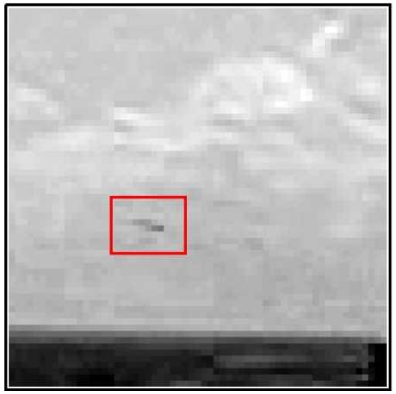
Frame 2893



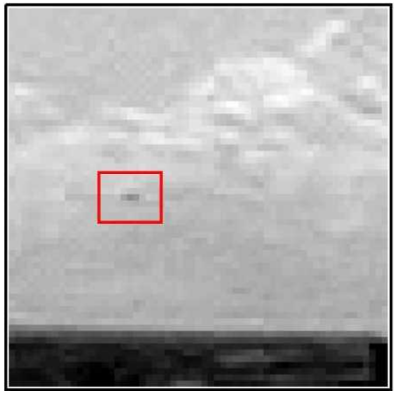
Frame 2894



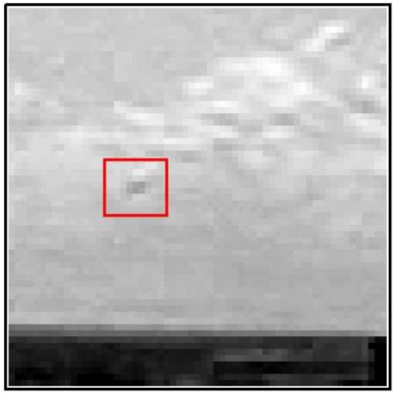
Frame 2895



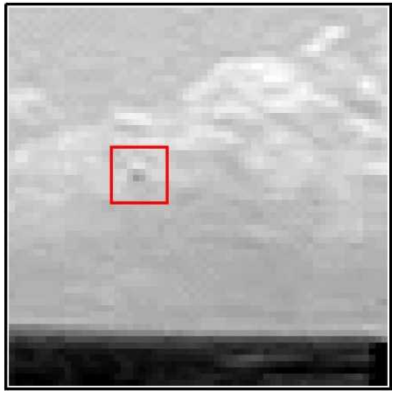
Frame 2896



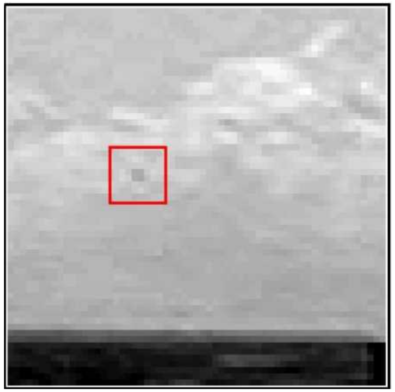
Frame 2897



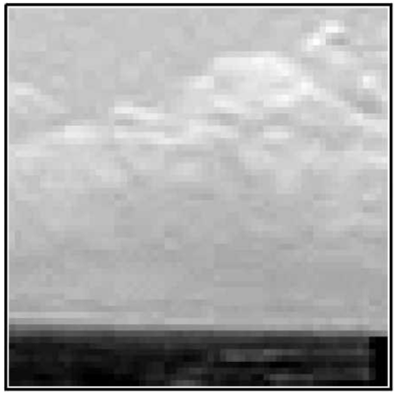
Frame 2898



Frame 2899



Frame 2900



A bird in flight is detected in seven frames.

Detection Parameters

DETECTOR PARAMETER:	*SETTING:*
FAPI Decay Rate	0.975
FAPI Decay Rate, Suppressed	0.99
Variance Decay Rate	0.99
Variance Decay Rate, Suppressed	1.0
Detection Threshold	6.0
Background Suppression Threshold	6.0
Variance Suppression Threshold	3.0
Jitter Standard Deviation	2.0

The same parameter values were used for the KAFB and border videos. However, a one-sided (positive deviations only) threshold was used for the infrared border data, while a two-sided (absolute value) threshold was applied for the visible KAFB data.