Overcoming Jitter Effects for Remote Staring Sensors (FY 08 – 09)

Sandia LDRD Day

14-Sep-2009 Albuquerque, NM

Principal Investigator: Katherine Simonson (5535)

Key Technical Staff: Tian Ma (5535), Ray Byrne (5535)

Project Managers: Bruce Malm (5531)
Mark Dowdican (5338)

Personnel Manager: John Feddema (5535)

Research Goal

- Our goal is the *automatic detection of small changes* in broad area surveillance.
  - We work with *low-resolution, staring, radiometric sensors*, which are subject to significant jitter.
  - Frame rates range from 10 – 30 Hz; algorithms must run in real time.
  - The magnitudes of the target changes are generally far smaller than the contrasts in the imaged scene, so targets cannot be detected by simply thresholding sensor frames.

- We have *worked extensively with real-world data* from a deployed, operational sensor.
  - As well as video sequences from a range of unclassified sources.

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.

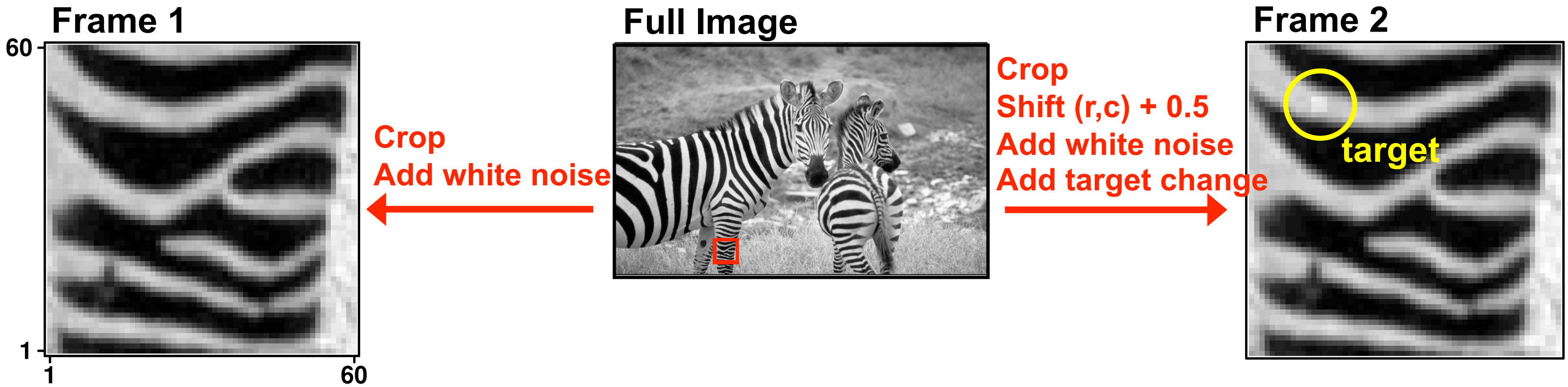
  - The background estimate may be a single prior frame or a more complex function evaluated over a window of recent frames.

- 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 change.

- Thus, *change detection in a high jitter environment is particularly challenging!*

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Mis-Registration

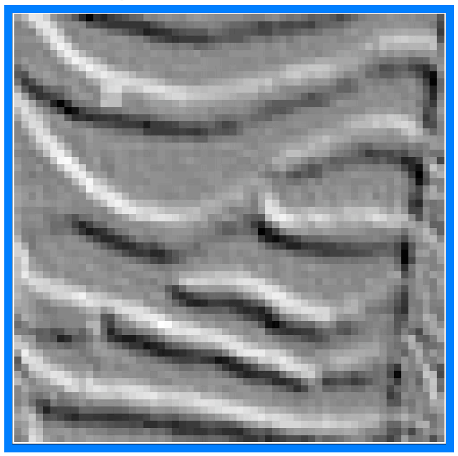


The difference between two frames slightly out of alignment is dominated by scene gradients larger than the target change.

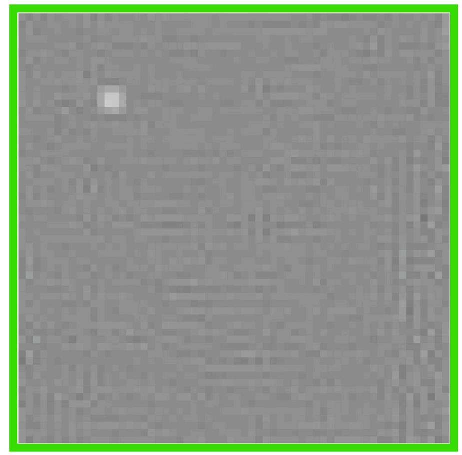
When the second frame is interpolated back into alignment with the first, the target signal is blurred but stands out readily.

The two difference frames are plotted in the same greyscale.

Frame 2 – Frame 1, Unregistered

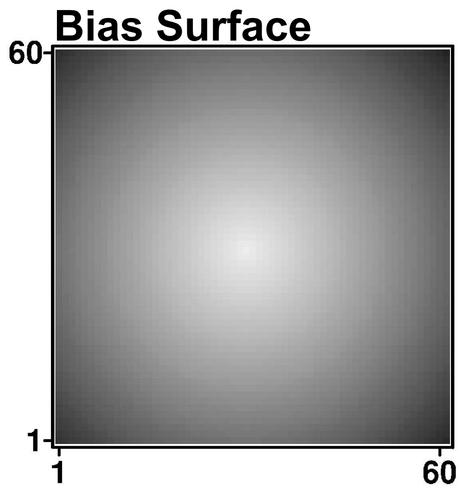


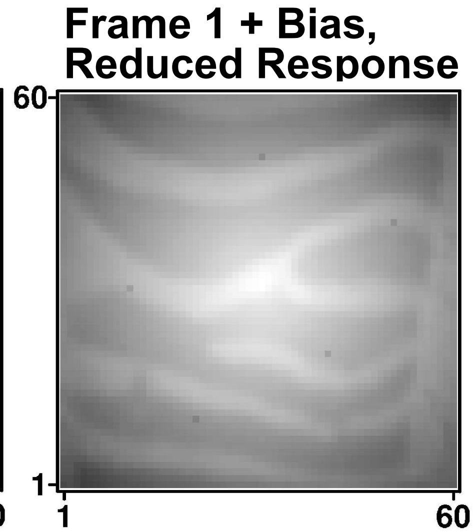
Frame 2 – Frame 1, Registered

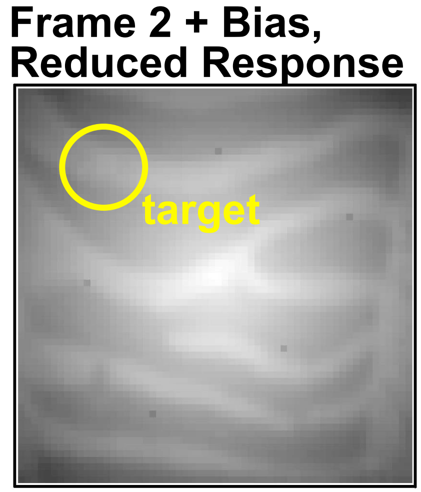


LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Sensor Artifacts

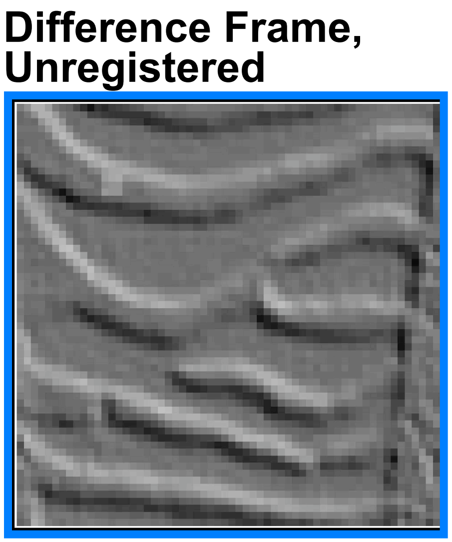


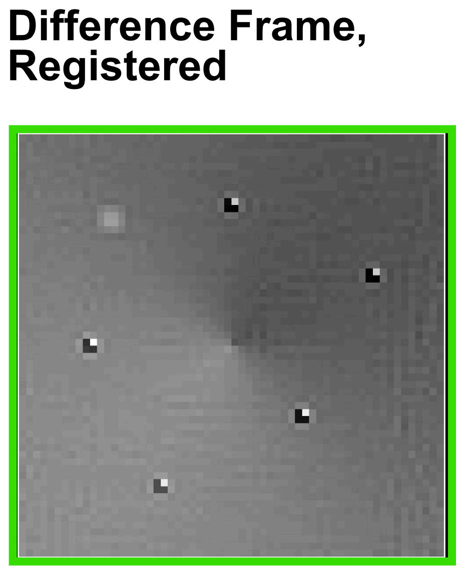




Artifacts in pixel space challenge solutions based on frame registration.

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





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 would have to be known and corrected for prior to scene registration.

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

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 generally not feasible at high frame rates (10 – 30 Hz) for low-quality data.**
  - **Even if jitter-induced offsets are known perfectly, all sensor artifacts (fixed pattern noise, self-emission, non-responsive pixels) have to be corrected for prior to frame transformation.**

- **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.**

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Normalized Differences

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

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

- Here, X represents the pixel's intensity at time t,
- B is the current background estimate,
- S is a current standard deviation estimate (based on data prior to t).

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

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

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 different standard deviation estimates.**

- **If min { | $Z_{SPATIAL}$|, | $Z_{TEMPORAL}$| } exceeds a fixed threshold, T1, the observed value of pixel (*k*,*h*) at time *t* is deemed inconsistent with both models, and a candidate change detection occurs.**

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

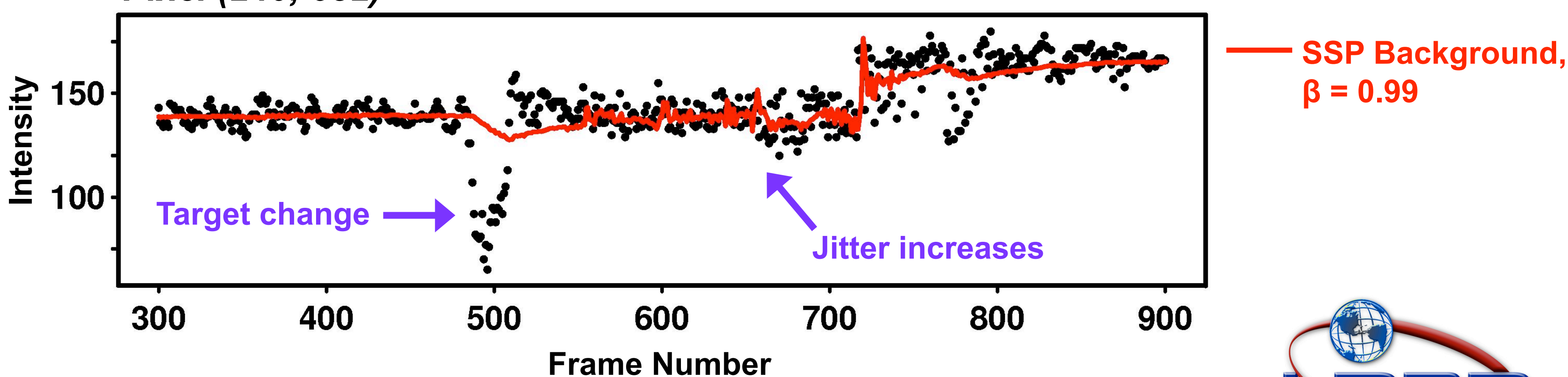
  - ***Area filtering:* Require detection in at least *N* connected pixels.**
  - ***Duration filtering:* Require detection in at least *M* consecutive frames.**

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Background Estimation

- ***Subspace projection (SSP)*** **is used to estimate the scene background.**
  - **- The general approach applied to the jitter problem has been around (at least) since Barry and Klop, 1983. Many algorithms exist for adaptive subspace estimation; some can update in real time.**
  - **- The goal is to capture the covariance structure of a sequence of frames in a low-dimensional, orthogonal subspace.**
  - **- Newly-observed frames are projected into this subspace; projection residuals are used to gauge change.**
  - **- To track gradual change in the scene (e.g., cloud motion), the subspace is updated after each frame. The decay rate, *β* ⊠ [0,1], is tunable.**

Pixel (240, 682)



LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Temporal Variances

- As the name implies, temporal variances are computed based on a recent time window of projection residuals. They are computed as follows:

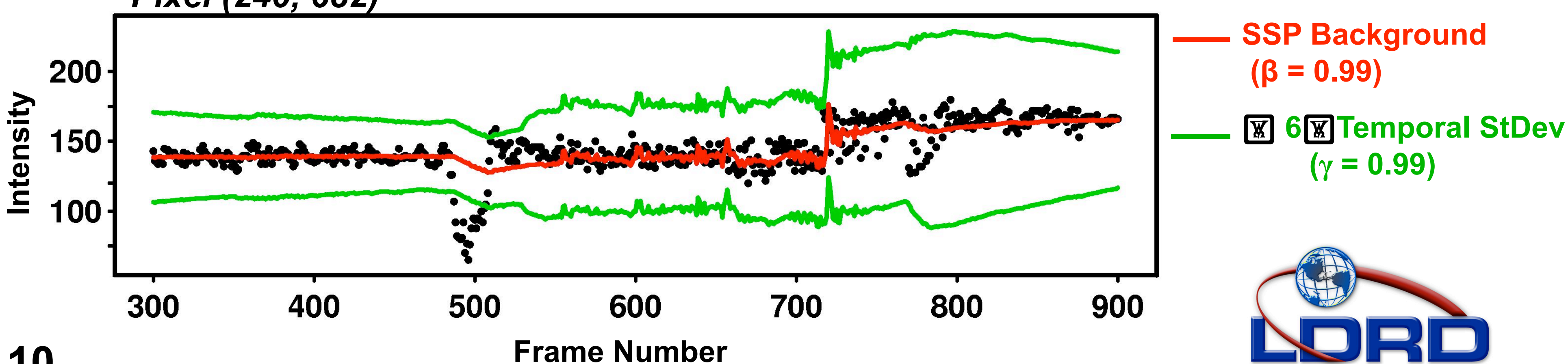
  1. Initialize with the sample variance over the first n frames, $V(k,h;n)$.
  2. For subsequent frames, update using:

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

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

  - The standard deviation estimate is not updated for any pixel (k,h) with normalized difference exceeding threshold T2.

Pixel (240, 682)



LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Spatial Estimation: Motivation

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

  - Used to normalize raw projection residuals.

- **If the jitter distribution is non-stationary, temporal estimates of pixel variance are inadequate: When jitter increases, large numbers of false alarms occur along scene gradients.**

  - Sub-space 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!***

- **This LDRD project has developed and matured a new mathematical concept for pixel variance estimation. The "spatial" approach can produce estimates that are robust to non-stationary jitter, based on a single frame.**

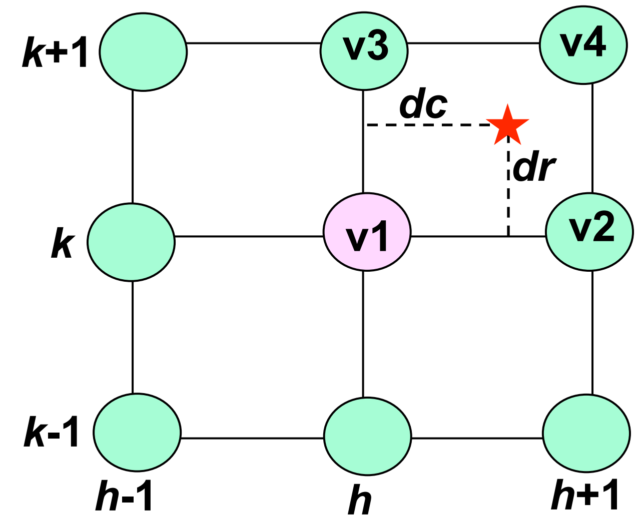
LABORATORY DIRECTED RESEARCH & DEVELOPMENT

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*:**

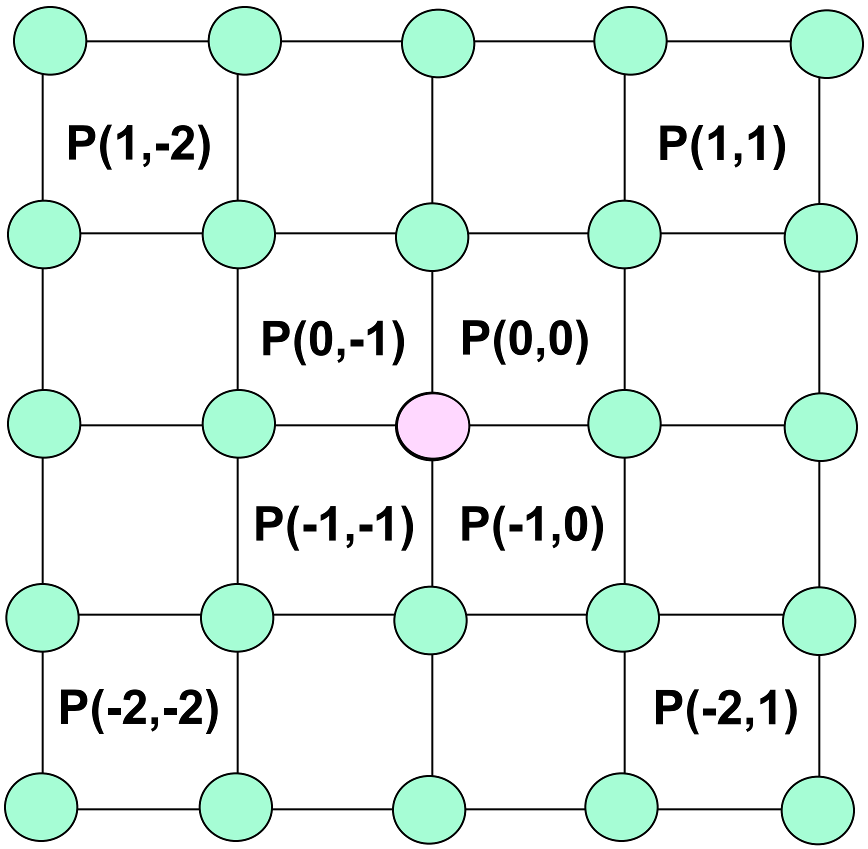


$$E(k,h;t) = v1 + dr(v3-v1) + dc(v2-v1) + dr \times dc(v1+v4-v2-v3)$$

- **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).**

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Conditional Expectation



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

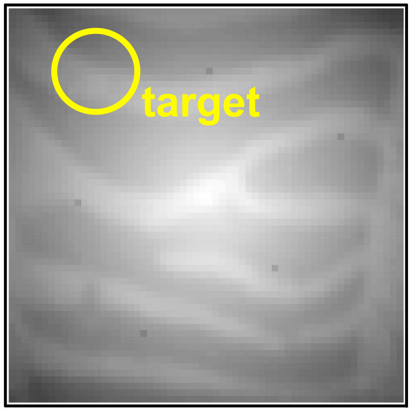
1) The probability of jittering into this cell at time t, and:
2) The expected pixel value (and its square) at t, given jitter into this cell.

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

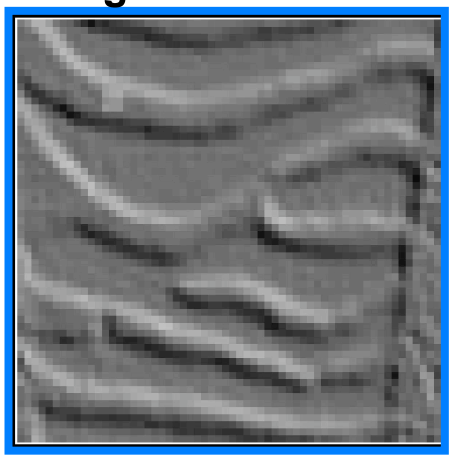
- **Estimates computed in this manner are *surprisingly robust to mis-specification of the jitter distribution*: They scale roughly linearly with the jitter standard deviation.**
  - **A good strategy is to set sigma conservatively (based on the worst jitter expected) and re-scale on a per-frame basis.**

Incorporating SSP and Spatial Variances

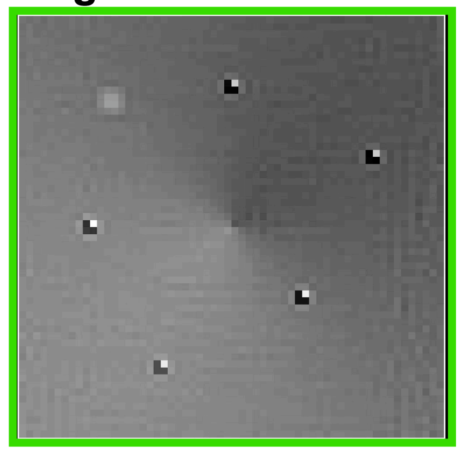
Frame 2 + Bias, Reduced Response



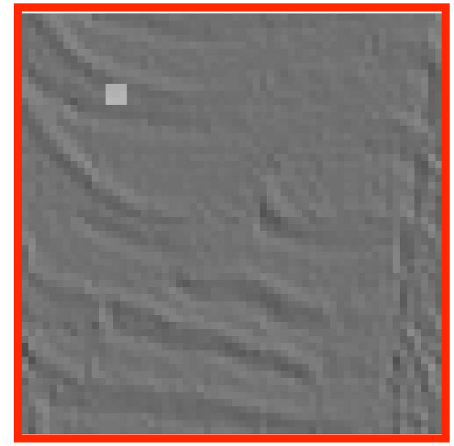
Difference Frame, Unregistered



Difference Frame, Registered



Raw SSP Residuals

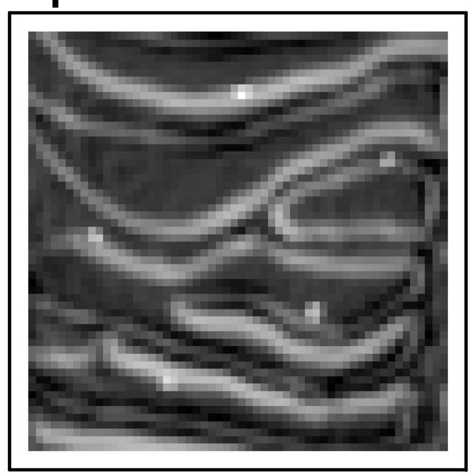


The principal subspace was estimated from 100 simulated jittered, noise-added versions of Frame 1 (with bias surface and reduced responsiveness).

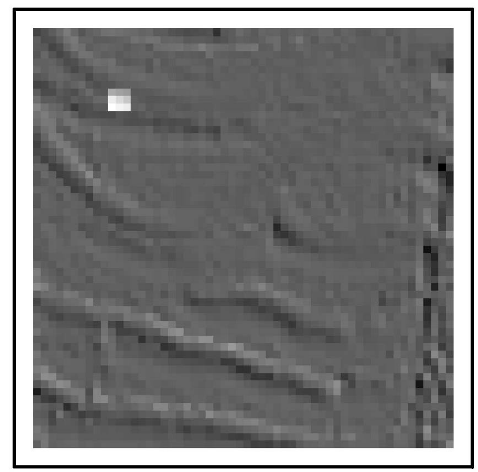
SSP residuals show less scene structure than the unregistered frame differences, and exhibit no sensor artifacts.

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



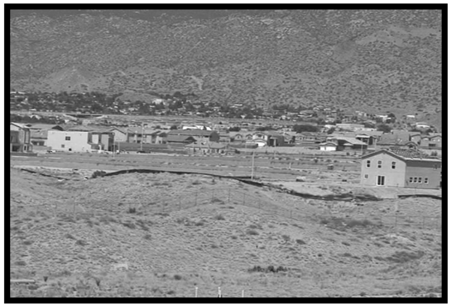
Normalized SSP Residuals



LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Video Processing Example

- **Jittery video data were collected by student interns from Sandia Dept. 6472 (Mobile Robotics).**
  - Brandon Cover, Robin Jones, Donald Kimmel.
  - Camera mounted on a fixed tower looked at several different scenes (KAFB, Four Hills Neighborhood)
  - Each sequence consists of 4000 frames (100 training, 3900 test).
  - Data are RGB but our method is currently limited to single-band (greyscale).
  - Extension to multi-band data is under study.



Parameter Settings:

$\beta = 0.99$	Background Estimation: Decay rate
$\gamma = 0.99$	Temporal Variances: Decay rate
$T2 = 3.0$	Temporal Variances: Suppression threshold
$\sigma = 4.0$	Spatial Variances: Jitter standard deviation
$N = 6$	Area filter: Minimum number of connected pixels
$T1 = 6.0$	Detection threshold

LDRD
LABORATORY DIRECTED RESEARCH & DEVELOPMENT

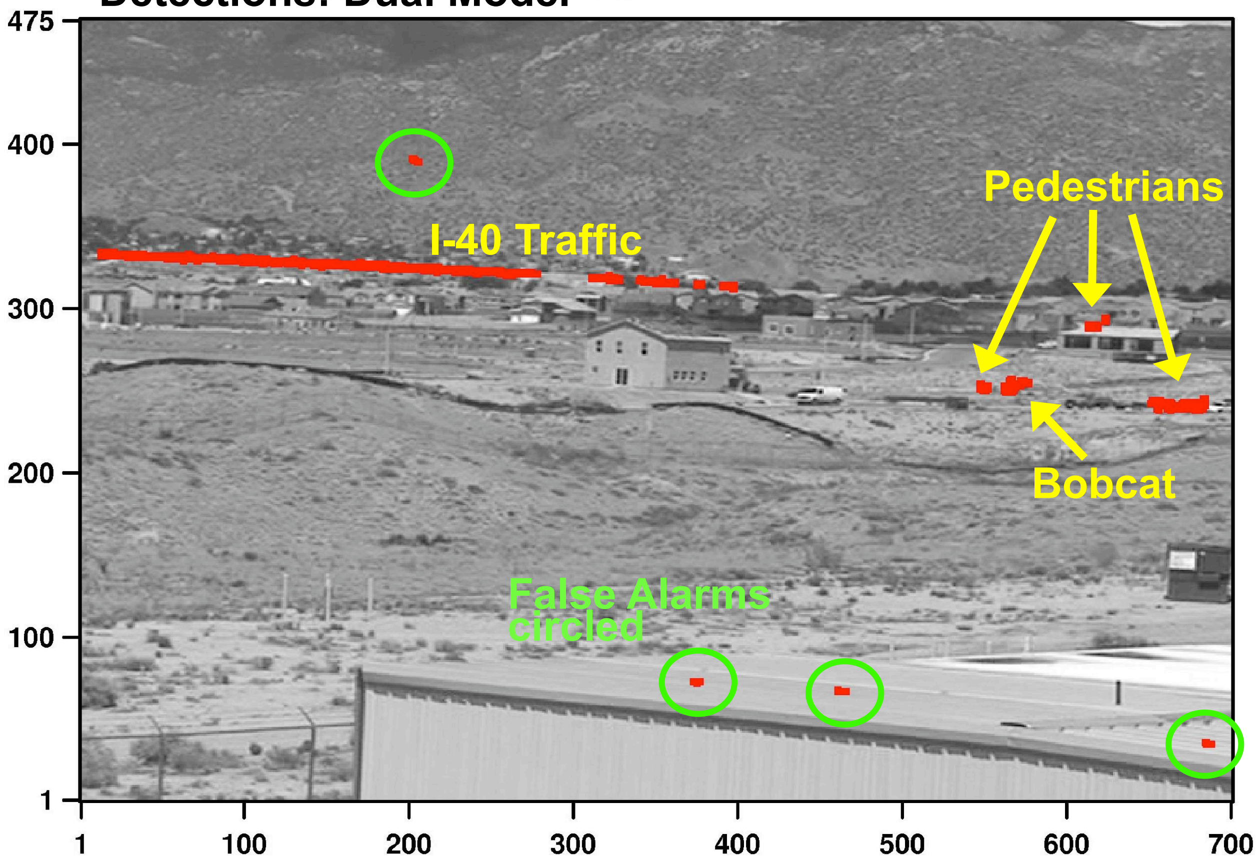
Run-Time Performance

- **The video frames are each 700 $\times$ 475 = 332,500 pixels.**
  - **With non-optimized research code, video processing runs at 6 frames per second (FPS).**
  - **For a sequence of smaller frames (200 $\times$ 380 = 76,000 pixels), a processing rate of 40 FPS can be achieved.**

- ***Parallel processing will enable real-time (30 Hz) performance for large frames.***
  - **Divide the frame into multiple blocks, overlapping along the edges.**
  - **Run background suppression and detection on each block in parallel.**
  - **Use appropriate logic to combine detections along block boundaries.**

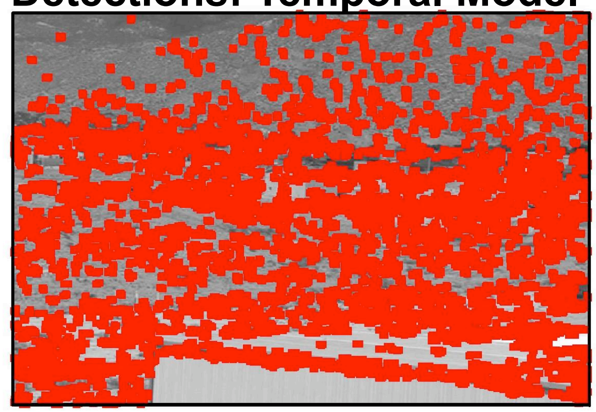
Summary Results: Frames 101 - 650

Detections: Dual Model



Detections: Spatial Model

Detections: Temporal Model



Jitter is non-stationary, but nominal (within 1-3 pixels) for the first 550 test frames. Many targets in motion are detected; some are very difficult to discern. Few false alarms occur with combined processing. Spatial-only detections include some noise-induced detections, while many false alarms caused by jitter occur for temporal-only detection.

Pixels shown in red had detection in *at least one* of the 550 frames.

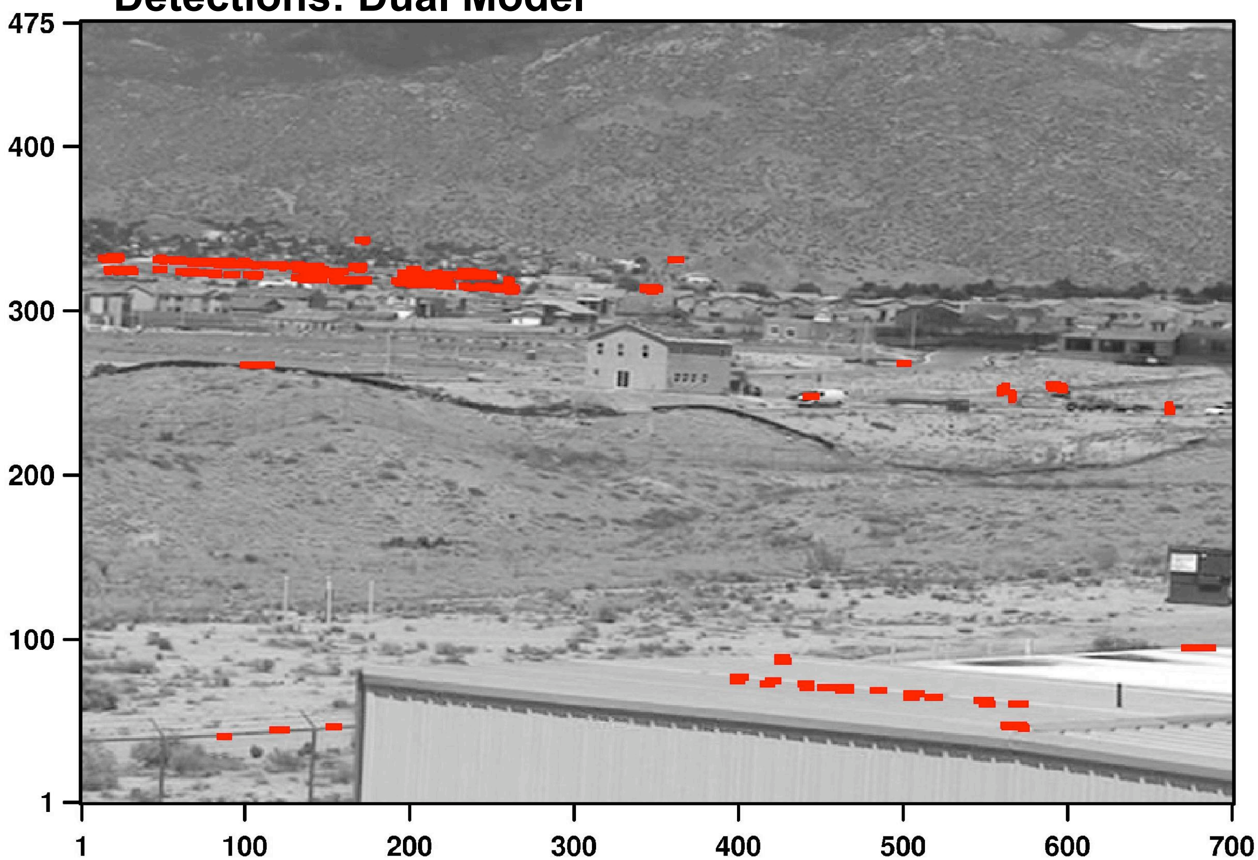
LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Detection Video: Frames 101 – 650, Nominal Jitter

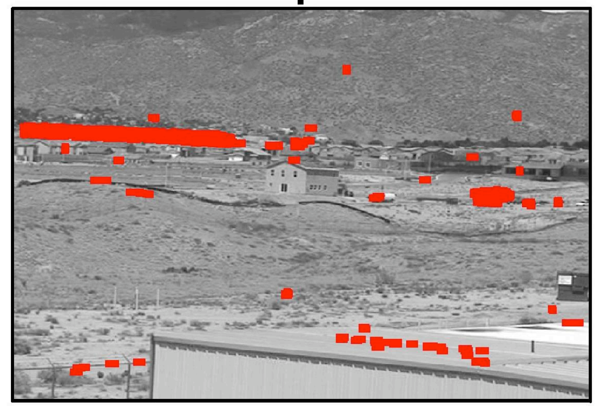
LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Summary Results: Frames 651 - 1200

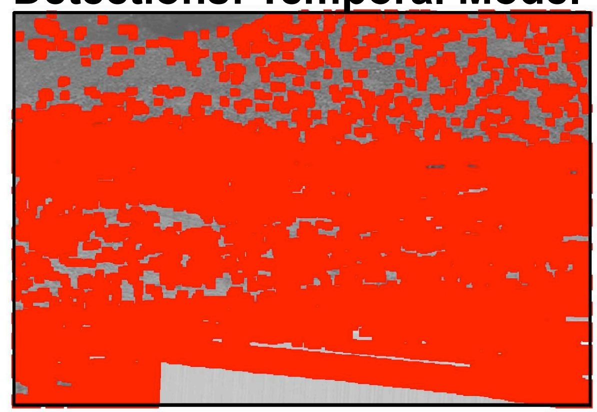
Detections: Dual Model



Detections: Spatial Model



Detections: Temporal Model



The jitter environment is more challenging (5+ pixels) for the next 550 frames. Target detection drops (largely due to increasing pixel standard deviations), and there are more false alarms. However, the system still finds many of the legitimate targets moving through the scene. Detection with temporal variances only fails completely.

Pixels shown in red have detection in *at least one* of the 550 frames.

LDRD
LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Detection Video: Frames 651 – 1200, High Jitter

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Summary and Path Forward

- **The change detection algorithms developed under this LDRD project show tremendous potential.**
  - **Spatial variance estimates enable robust change detection, even in challenging jitter environments.**
  - **A patent application has been filed (Simonson and Ma).**
  - ***A two orders of magnitude reduction in the false alarm rate* (at fixed target detection percentage) has been achieved for challenging data from an operational sensor.**
  - ***Initial transition expected within a year*.**

- **Work will continue under funding from several sources:**
  - **Development, test, and integration of algorithm into baseline operational software under external sponsorship.**
  - **FY2010 – 12 LDRD project on improved tracking of closely-spaced and maneuvering targets (Principal Investigator Tian Ma)**

- ***More opportunities sought* for current or future systems!**

LDRD
LABORATORY DIRECTED RESEARCH & DEVELOPMENT