A Comparison of Background Subtraction Algorithms for Detecting Avian Nesting Events in Uncontrolled Outdoor Video

Kyle Goehner, **Travis Desell**

Department of Computer Science
University of North Dakota

Rebecca Eckroad, Leila Mohsenian, Paul Burr, Nicholas Caswell, Alicia Andes, and Susan Ellis-Felege

Department of Biology
University of North Dakota

September 1, 2015 eScience 2015, München, Germany

Wildlife@Home

http://csgrid.org/csg/wildlife

What is Wildlife@Home?

- A *citizen science* project that combines both crowd sourcing and volunteer computing.
- Users volunteer their brain power by observing videos and reporting observations.
- Users volunteer their computer power by downloading videos and performing.
- A scientific web portal to robustly analyze and compare results from users, experts and the computer vision techniques.

Between 2012 and now, Dr. Ellis-Felege has gathered over 85,000 hours of avian nesting video from the following species:

1. Sharp-tailed grouse (*Tympanuchus phasianellus*), an important game bird and wildlife health indicator species.
2. Piping plovers (*Charadrius melodus*), a federally listed threatened species.
3. Interior least terns (*Sternula antillarum*), a federally listed endangered species.

A recent collaboration with Ducks Unlimited added another 15,000 hours of Blue Winged Teal (*Anas discors*) nesting video.

We have also recently received over 2 million motion sensor camera images and ~100,000 aerial images taken by UAVs from a new Hudson Bay project.

Sharp-tailed Grouse

Piping Plover

The three species (Grouse, Plover and Tern) investigated in this work are ground nesting birds.

Sharp-tailed grouse nest in the dense grass (top left). Nests were monitored in areas of high oil development, moderate oil development and no oil development (protected state land).

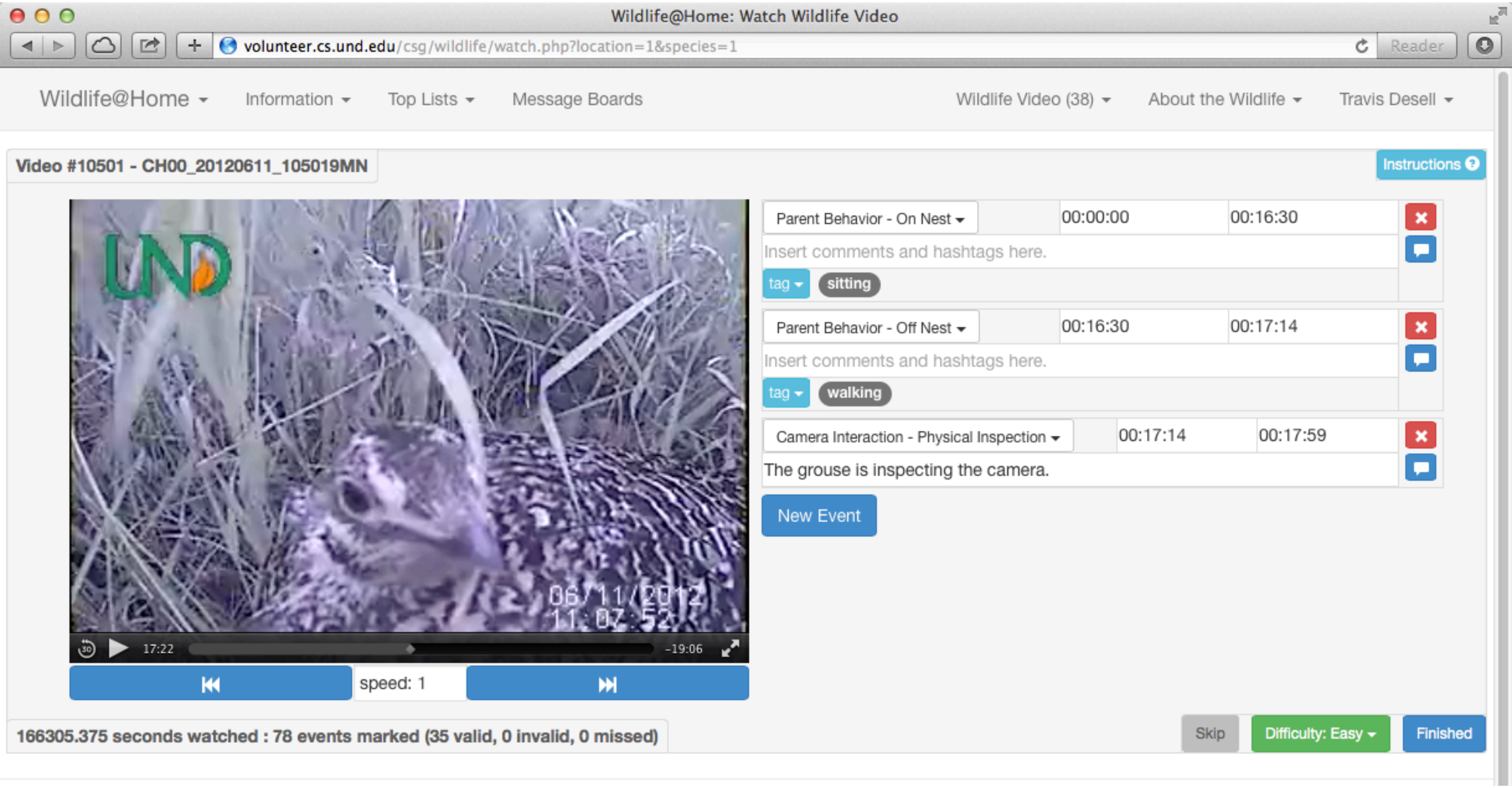
Piping plover and interior least tern are shore nesting species (top right). Nests were monitored along the Missouri River in North Dakota.

Most grouse video is sleeping birds and grass blowing in the wind.
But occasionally, interesting things happen.

Piping plover and tern video is more interesting, with active bi-parental involvement and less obscuring vegetation.

There are many challenges:

1. Dramatically changing weather conditions
2. Dawn/Day/Dusk/Night lighting conditions
3. Model species (sharp tailed grouse and piping plover) and some predators have cryptic coloration (camouflage).
4. Moving vegetation and insects can cause false negatives.
5. Lower quality video due to limitations on cameras.



We have been information about the video through a crowd sourcing interface, and a similar interface used by research assistants in biology.

Background Subtraction Methods

Mixture of Gaussians

MOG describes the probability of a pixel belonging to the background as a sum of Gaussians:

$$f_{\mathbf{X}}(X|\Phi) = \sum_{k=1}^{K} P(k) \cdot f_{\mathbf{X}|k}(X|k,\theta_k)$$

Where *P(k)* is the probability of the surface *k* appearing in the pixel view, and $f_{X|k}$ is the Gaussian distribution for surface k with Phi being the set of theta input parameters for the Gaussian distributions describing each feature.

P(k), uk, and thetak can be estimated with running averages calculated at each frame, and $f_{X|k}$ can be estimated by a boolean value which is true for a pixel value if it is within 2.5 standard deviations of the mean.

ViBe

Vibe stores the history of 20 previous pixel values, and compares new values to this pixel history.

If a pixel is within some threshold of any pixel within this stored model, it is classified as background.

The background model is updated stochastically, with each new pixel value having a 1/16 chance to replace one of the 20 stored pixel values selected at random. If a replacement is done, there is an additional 1/16 chance of also updating one randomly selected neighborhood pixel's previous values.

Pixel-Based Adaptive Segmentation (PBAS)

PBAS is an extended version of ViBe which adjusts the threshold for selecting a pixel as background dynamically.

This is done using another set of 20 values, however in this case these are the minimal decision distance (minimum distance between an updated pixel and the previous 20 pixels). The average of these 20 minimum decision values is used to calculate the threshold, $R(x_i)$, which increases/decreases by a user defined scale whenever it is above or below that average.

Motion Detection for Avian Nesting Video

ViBe and PBAS were modified and compared to MOG for this work:

1. They were made 2nd frame ready - the initial 20 previous pixels were selected at random from the first image.

2. An open/close filter was added to reduce foreground detection noise. This essentially smoothes the image, aiding in the reduction of video artifacts.

3. The convex hull of any connected foreground features used as foreground mask. This increases the selected foreground area, as in many cases the head and other parts of the bird are foreground while the rest of the bird matches the background too well due to cryptic coloration.

Motion Detection for Avian Nesting Video

With these additions, the foreground mask needed to be converted to a measure of the probability of an event of interest occurring.

The count of foreground pixels is used as a time series of data points, which is smoothed by an exponential moving average:

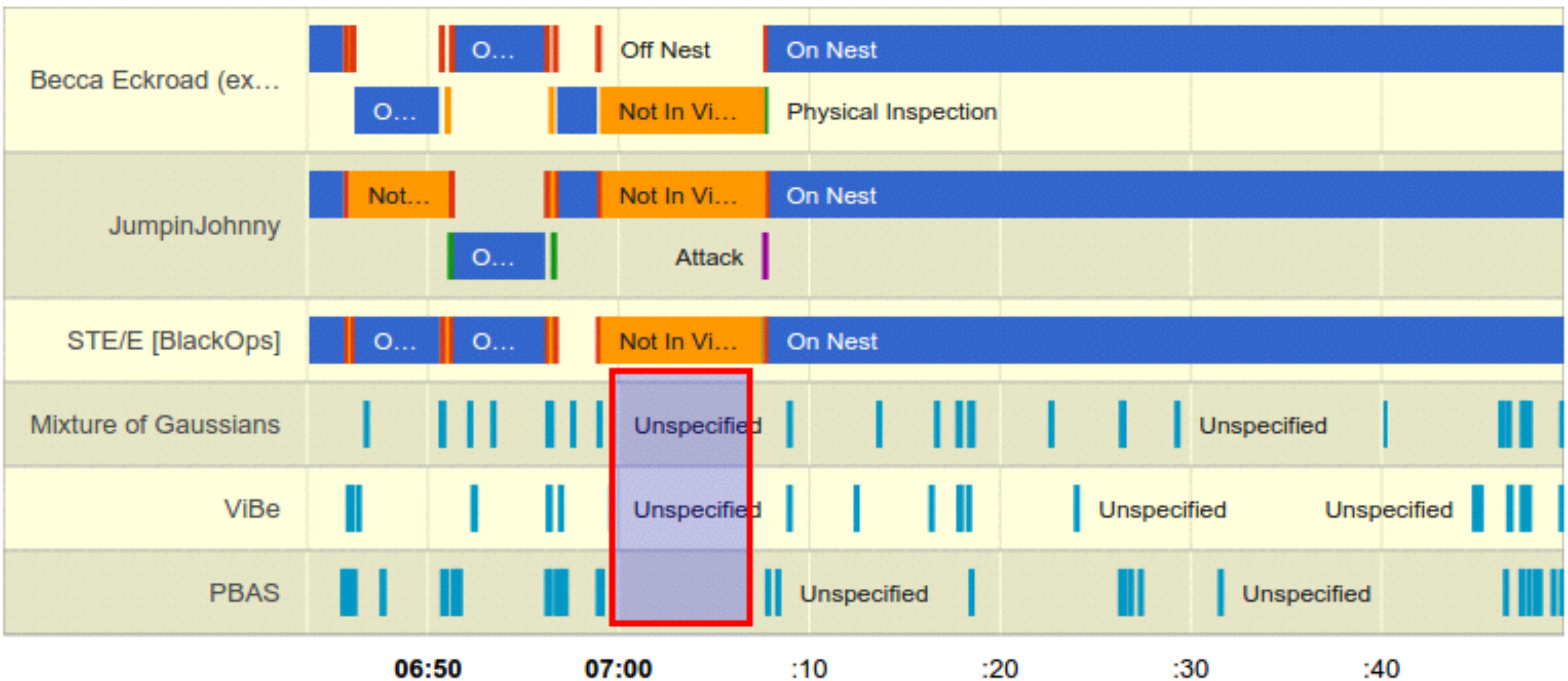
$$m_t = \alpha \cdot x_t + (1 - \alpha) \cdot m_{t-1}$$

Where m_t is the mean at time unit t, x_t is the number of foreground pixels at time t, and alpha is the learning rate.

If at any time, x_t is greater than the three standard deviations from the time series mean, m_t, then that frame is flagged as having an event.

Motion Detection for Avian Nesting Video

This can then be used to generate time series which can be compared to crowd sourced and expert video observations:



Results

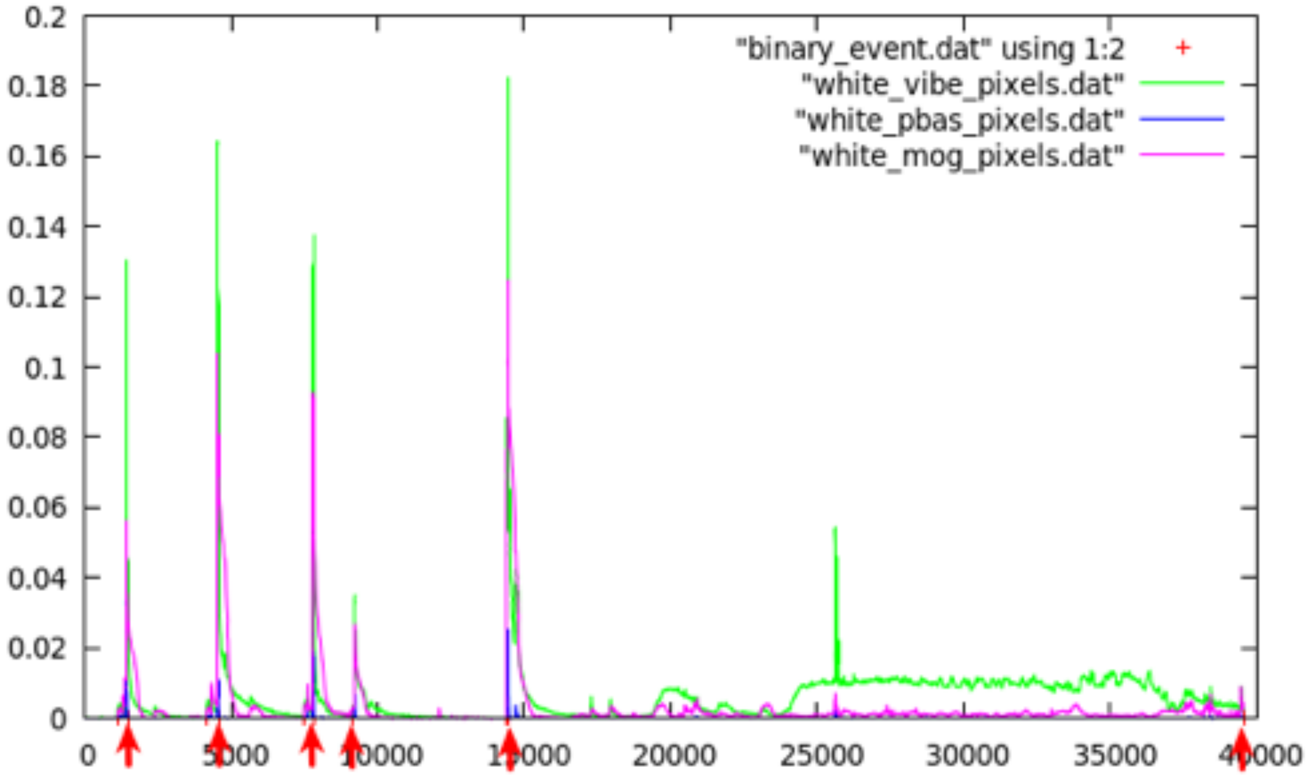
Experiments

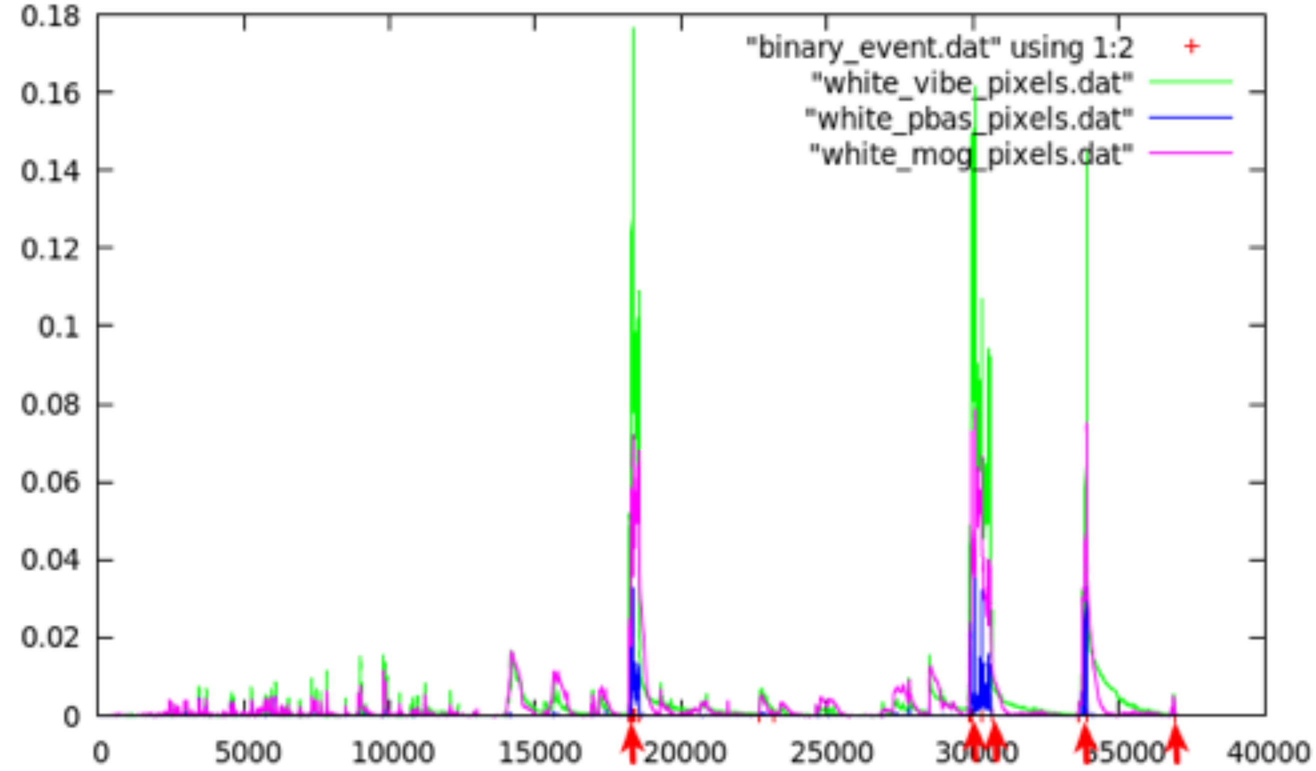
MOG, as well as our modified ViBe and PBAS were run over 105 tern and plover videos (77.05 total hours), and 109 sharptailed grouse videos (205.39 total hours).

Video lengths range from 30 minutes to 2 hours, and each algorithm ran at ~10 frames per second.

Results were gathered using a Mac Pro with 12 logical cores, and took approximately 48 hours.

Detecting Interesting Events





The above shows two time series of detected foreground pixels. Red arrows at the bottom show the beginning and end of scientist observed events.

The algorithms were described as correctly detecting an event if it fell between the start and end time of a user observed event.

Detecting Interesting Events

TABLE III
ALGORITHM ACCURACY VS EXPERT SCIENTISTS ON GROUSE NESTS

Event Type	Event Count	MOG	ViBe	PBAS
Not In Video	284	274	258	270
Eggshell Removal	6	4	5	5
In Video	130	128	129	129
Predator	6	5	5	5
Unspecified	2	2	2	2
Attack	2	2	2	2
Physical Inspection	60	52	56	56
Observation	44	41	39	41
On Nest	216	196	174	178
Off Nest	492	470	439	461

TABLE I
ALGORITHM ACCURACY VS EXPERT SCIENTISTS ON TERN AND PLOVER NESTS

Event Type	Event Count	MOG	ViBe	PBAS
Preen	180	170	138	147
Scratch	4	4	2	2
Not In Video	732	632	578	607
Nest Exchange	22	16	16	16
Foraging	82	71	52	56
Adult-to-Adult Feed	20	6	6	6
Nest Defense	4	4	4	4
Predator	12	10	7	9
Non-Predator Animal	22	16	15	15
Unspecified	350	93	66	78
On Nest	932	665	582	608
Off Nest	2312	1960	1775	1876

The above charts show how well each algorithm matched up to events classified by project scientists (the paper also includes comparisons to our citizen scientists). All the algorithms performed well detecting events, with MOG detecting the most.

Detecting Interesting Events

TABLE V
ALGORITHM ACCURACY WITH CONSENSUS VS EXPERT SCIENTISTS ON TERN AND PLOVER NESTS

Event Type	Event Count	Any Alg	All Alg	MOG & ViBe	MOG & PBAS	ViBe & PBAS
Preen	180	174	137	138	143	137
Scratch	4	4	2	2	2	2
Not In Video	732	635	576	576	606	576
Nest Exchange	22	16	16	16	16	16
Foraging	82	73	51	52	54	51
Adult-to-Adult Feed	20	6	6	6	6	6
Human	2	0	0	0	0	0
Nest Defense	4	4	4	4	4	4
Predator	12	11	6	6	8	7
Non-Predator Animal	22	19	12	12	14	13
Unspecified	350	94	66	66	77	66
On Nest	932	669	572	580	606	572
Off Nest	2312	1974	1763	1769	1868	1763

The above chart shows results for combining the different algorithms. Having a consensus from multiple algorithms tended to lower event detection.

Analysis of False Positives

TABLE VI
ALGORITHM FALSE POSITIVES VS EXPERT SCIENTISTS

	MOG		ViBe		PBAS	
Species	μ	σ	μ	σ	μ	σ
Grouse	139.67	144.76	74.31	95.92	73.83	100.64
Tern	5.78	35.37	2.76	15.86	1.58	6.89
Plover	4	7.63	0.50	1.07	0.63	1.41

TABLE VII
ALGORITHM FALSE POSITIVES VS CITIZEN SCIENTISTS

	MOG		ViBe		PBAS	
Species	μ	σ	μ	σ	μ	σ
Grouse	118.27	136.17	53.14	74.65	53.90	82.10
Tern	0.41	1.74	0.22	0.80	0.15	0.46

An analysis of false positives was provided. A false positive was measured as the number of events classified during a user classified *Not In Video* event.

The grouse video, which has significant amounts of high wind and moving vegetation had far more false positives (as to be expected). On the other hand it also had a very high standard deviation - suggesting that for videos without high wind and moving vegetation the background subtraction performed well.

Plover and Tern video had significantly less false positives, however the standard deviation was high, suggesting that for some videos (high wind or light fluctuations) these algorithms performed poorly.

While MOG detected the most events, it also had significantly more false positives.

Effectiveness of Background Subtraction

The modified PBAS and ViBe both performed well in detecting events, while MOG had rates of false positives that were too high to be effective.

While PBAS and ViBe were highly effective for a large number of video, there still remains a challenging subset of video with high wind and/or frequent lighting changes which will require more advanced techniques.

What's Next?

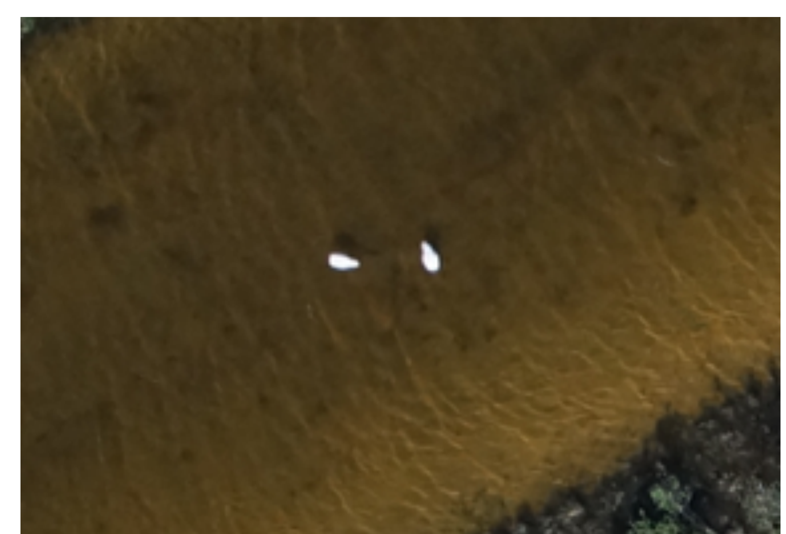
We have used Wildlife@Home's volunteered computers to run the motion detection methods over all the collected video. Results have been incorporated as a timeline into the user interface. Users can click on the timeline to skip ahead to areas of interest.

What's Next?

The motion detection methods used, especially ViBe and PBAS work well on "easy" segments of the video.

New methods need to be developed to handle the challenging sections of video with rapidly changing light conditions and/or windy rapidly moving vegetation. Potential ideas: convolutional neural networks, Retinex to normalize brightness.

Expanding crowd sourcing to imagery from UAVs and motion sensing cameras taken in North Dakota and the Hudson Bay, Canada.



Reproducibility

All the videos and observations used in this work have been made available in the first Wildlife@Home data release:

http://csgrid.org/csg/wildlife/data_releases.php

Video ID	Species	Download	Event	Start Time (s)	End Time (s)	Tags	Comments
10720	1	[mp4] [ogv]	Parent Behavior - On Nest	5968.2	7705	#sitting	
			Parent Behavior - Off Nest	5962.600689	5968.2	#walking	
			Parent Behavior - Not In Video	3942.800001	5962.600689		
			Parent Behavior - On Nest	0	3936.100072	#sitting	
			Parent Behavior - Off Nest	3936.100072	3942.800001	#standing#walking#flying	
10724	1	[mp4] [ogv]	Parent Behavior - Off Nest	1593	2700		
			Parent Behavior - Not In Video	1298	1593		
			Parent Behavior - Off Nest	1255	1298		
			Parent Behavior - Not In Video	0	1255		
10860	1	[mp4] [ogv]	Parent Behavior - On Nest	0	3144	#sitting	
			Parent Behavior - Off Nest	3144.753429	3145.385429	#flying	

And all Wildlife@Home source code is freely available on GitHub:

https://github.com/travisdesell/wildlife_at_home

Acknowledgements

Wildlife@Home is currently being supported by NSF award no. 1319700 through the Division of Intelligent Information Systems's Information Integration and Informatics (III) program.

Wildlife@Home has been generously supported by a collaborative research award and new faculty SEED grant from UND's Office of Research Development and Compliance. The project's video streaming server is hosted by UND's Computational Research Center and the volunteer computing server is hosted by UND's Scientific Computing Center. DNA@Home is under partial support from a Basic Sciences SEED Grant.

North Dakota Game and Fish has provided financial support for field logistics to collect sharp-tailed grouse videos.

The US Geological Survey has provided financial support for camera equipment, video storage, and field assistance to collect data for the piping plover and interior least tern.

And of course all our volunteers.

Thanks!

Questions?

http://people.cs.und.edu/~tdesell/

http://csgrid.org

tdesell@cs.und.edu