PENNSTATE

ImageScan

A target detection processing chain using ArcMap

Peter Telek

Jun 2012

Advisor: Jay Parrish

PENNSTATE

Outline

- Background
- Overview
- Goals and Objectives
- Proposed Methodology
- Project Timeline
- Anticipated Results
- Summary
- List of Acronyms
- References

Background

- Increase in our ability to gather (intel) data
  - Especially geospatial data
  - Collecting more than 1500% more data than 5 years ago
  - ISR task force system can gather more than 53 TBytes data/day

Too Much Data?

Can you ever have too much data?

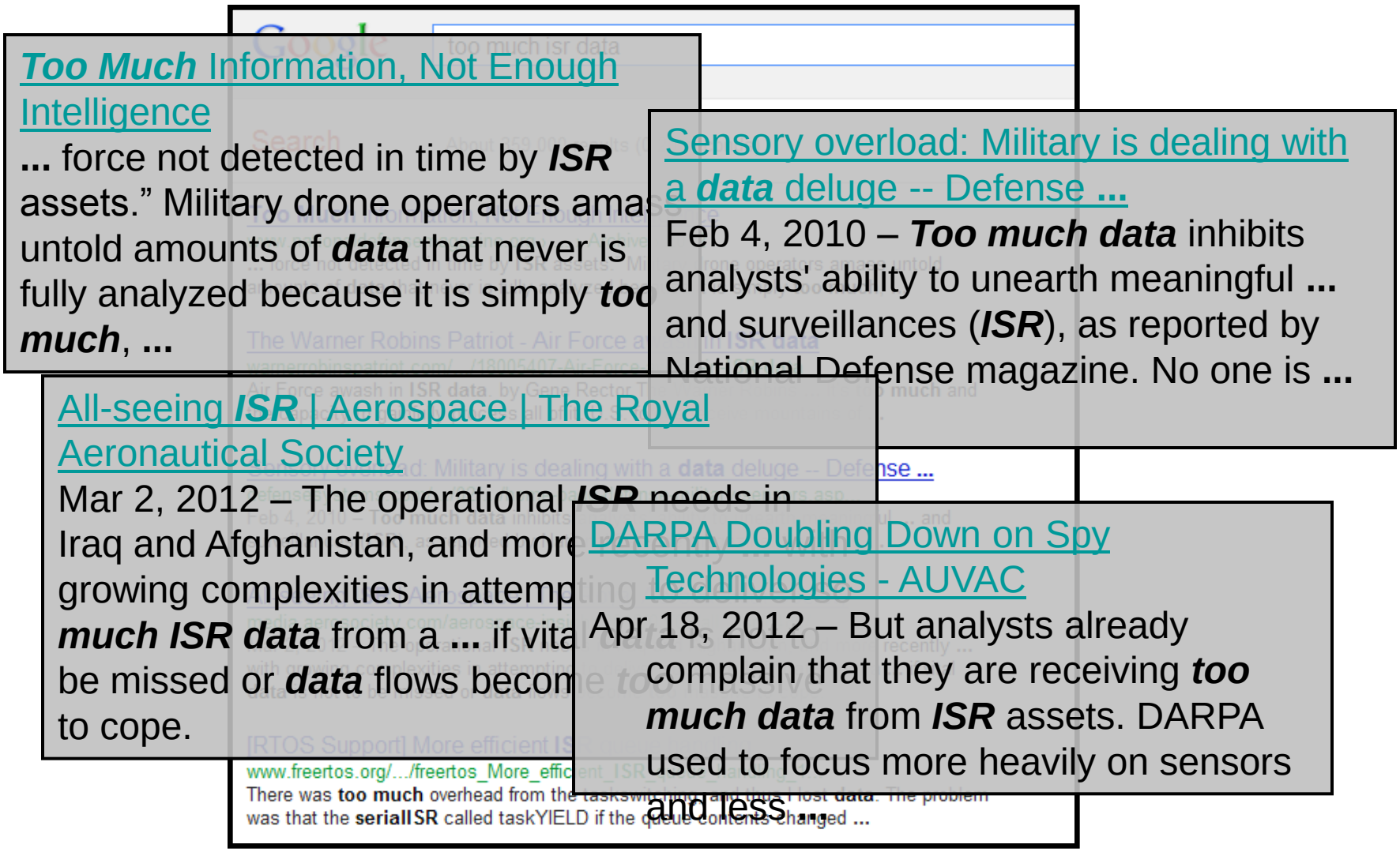
A simple search shows …

PENNSTATE

Too Much Data?

PENNSTATE

Too Much Data?



Too Much Information, Not Enough Intelligence

... force not detected in time by ***ISR*** assets." Military drone operators amass untold amounts of ***data*** that never is fully analyzed because it is simply ***too much***, **...**

Sensory overload: Military is dealing with a ***data*** deluge -- Defense **...**

Feb 4, 2010 – ***Too much data*** inhibits analysts' ability to unearth meaningful **...** and surveillances (***ISR***), as reported by National Defense magazine. No one is **...**

All-seeing ***ISR*** | Aerospace | The Royal Aeronautical Society

Mar 2, 2012 – The operational ***ISR*** needs in Iraq and Afghanistan, and more growing complexities in attempt ***much ISR data*** from a **...** if vita be missed or ***data*** flows becom to cope.

DARPA Doubling Down on Spy Technologies - AUVAC

Apr 18, 2012 – But analysts already complain that they are receiving ***too much data*** from ***ISR*** assets. DARPA used to focus more heavily on sensors and less **...**

PENNSTATE

Background (cont.)

- Increase in our ability to gather (intel) data
  - Especially geospatial data
  - Collecting more than 1500% more data than 5 years ago
  - ISR task force system can gather more than 53 TBytes data/day
- Real challenge is to improve our Processing, Exploitation/Analysis, and Dissemination (PED)
  - Continuous goal is to improve PED efficiency
  - Change/improve data dissemination
  - Data transfer - Identify data transmission bottle necks
  - Maximize our resource efficiencies
    - Identify high priority data for detailed analysis
    - Automate routine analysis

Background (cont.)

- Automation is a must
  - Computers are cost effective tools
  - Complete repetitive tasks
  - Analyst's time is expensive/valuable – manage it wisely
  - Need to build processes to …
    - maximize their efficiency by reducing their time spent on data searches and preprocessing
    - help prioritize their efforts
      - ID key data sets
      - ID points of interest
  - Catalog Results are critical with ISR applications, distributed/federated analysis, and information sharing
    - Build historical databases
    - Useful in trend analysis
    - Useful in detecting anomalies
      - Focus analyst on unusual (potentially) key information

Overview

- The goal of this proposal is to:
  - Create an automated or semi-automated image scanning and reporting processing chain (using ArcMap and other software if needed)
  - Alert customers, analysts, or end-users with a geospatial report of target locations and a subset list of priority images to analyze

Goals and Objectives

- ImageScan will create a system to
  - process data in an (ingest) directory,
  - read/enter a set of criteria [bounding box(es),
  - start stop dates,
  - detection threshold limits (size and intensity)],
  - summarize the data (footprints, collection date/time),
  - identify/highlight data sets that meet the criteria for analyst review; and
  - create reports (web pages, thumbnails, dynamic pages), and
  - create shapefiles.
    - The shapefiles would be used to update a geodatabase for treading or other sensor integration/fusion analysis efforts.
- *Caveat: This is meant to be a robust scanning technique to help pick out targets of interest, data sets of interest with a unique (I think) method of detecting and bounding high contrast targets. I have no doubt this can be improved upon. Additional work would be needed to make any part of this project more efficient.*

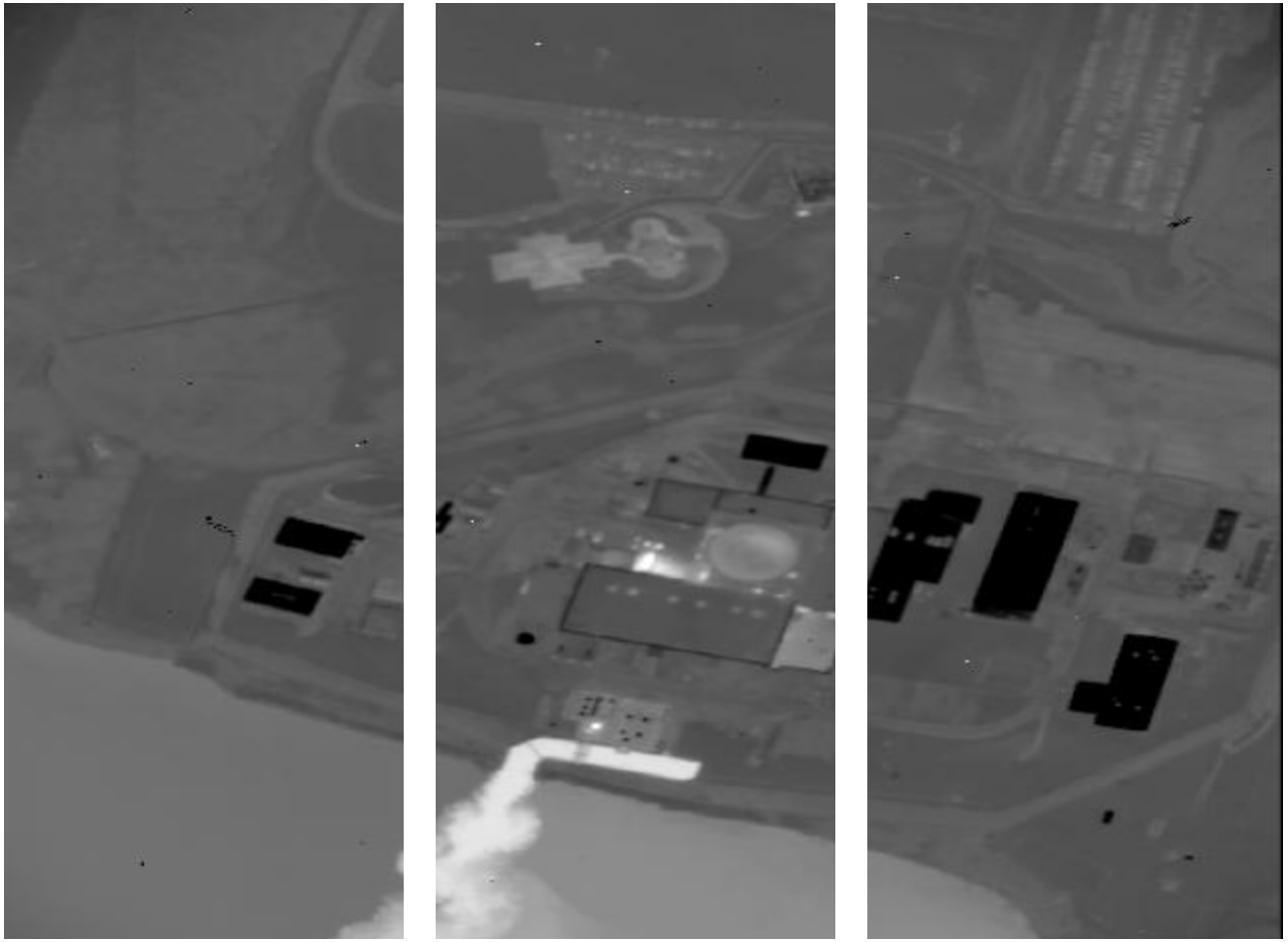
PENNSTATE

The Objective

- **The Problem**: With a library with over a thousand raster images, identify all data sets with high intensity values pixels (that exceed DN xxxx).
- For each raster image:
  - Determine the location of each detection
  - Do not report every pixel location; Cluster neighboring pixel (detections) together so that a multiple pixels target is reported as one target detection. For example, an oval 4x10 pixel target is reported as 1 target location and 1 polygon and not 40 individual target detections.
  - Locate the center of each target detection
  - Create a polygon for each target detection
  - Bonus: eliminate false alarms - in this case, small targets - single pixel detections
  - Create a library of detections that can be searched geographically, temporally
  - Repeat for the next image
- Repeat this every week (or every data delivery cycle)
- The results from ImageScan will be integrated with other data sources for a separate research project which is outside the scope of this capstone
- **This Capstone project will create a working prototype that will work on a few images.**

PENNSTATE

Sample Images

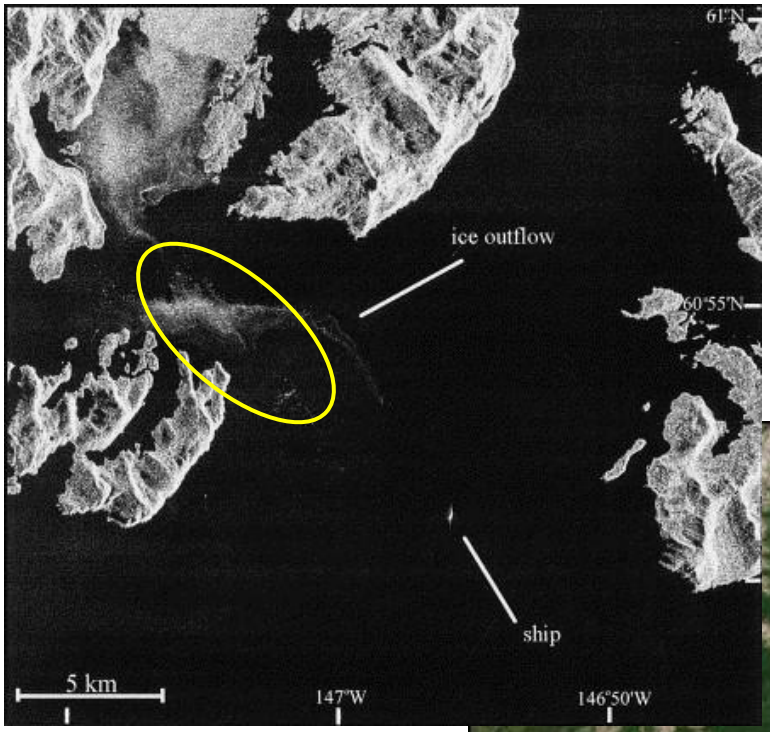


WASP Thermal MWIR Images of Ginna Power Plant

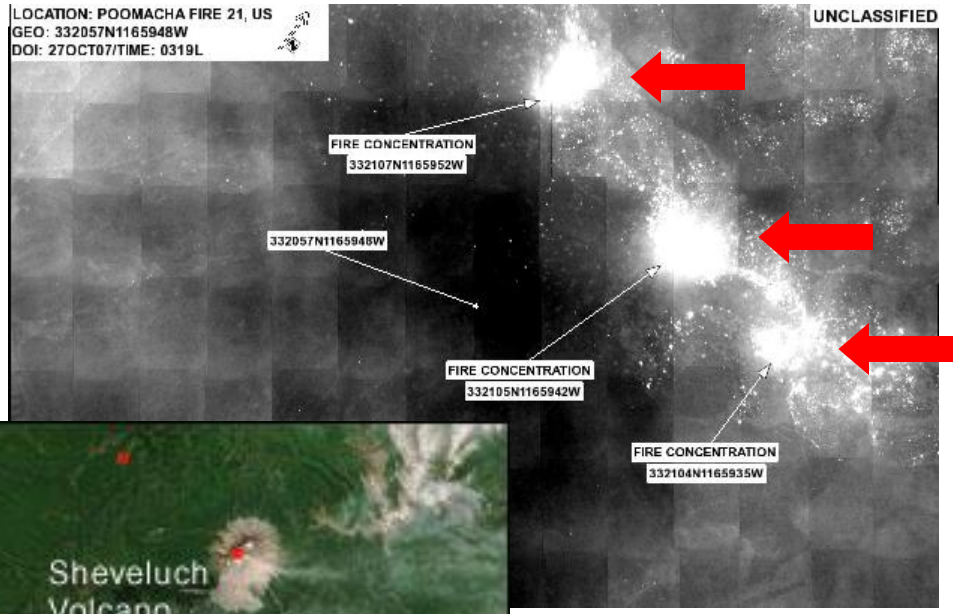
http://www.google.com/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=4&ved=0CDUQFjAD&url=http%3A%2F%2Flandsat.usgs.gov%2Fdocuments%2F6a_Schott_Overview_of_DIRS_Research%2520Activities.pdf&ei=N33JTqLfNa7y2gW6q4ngDw&usg=AFQjCNHKb35F1bXIjMoikF9xyUvBxEInCQ&sig2=vop2MSeAqS7z707GnHjOaw

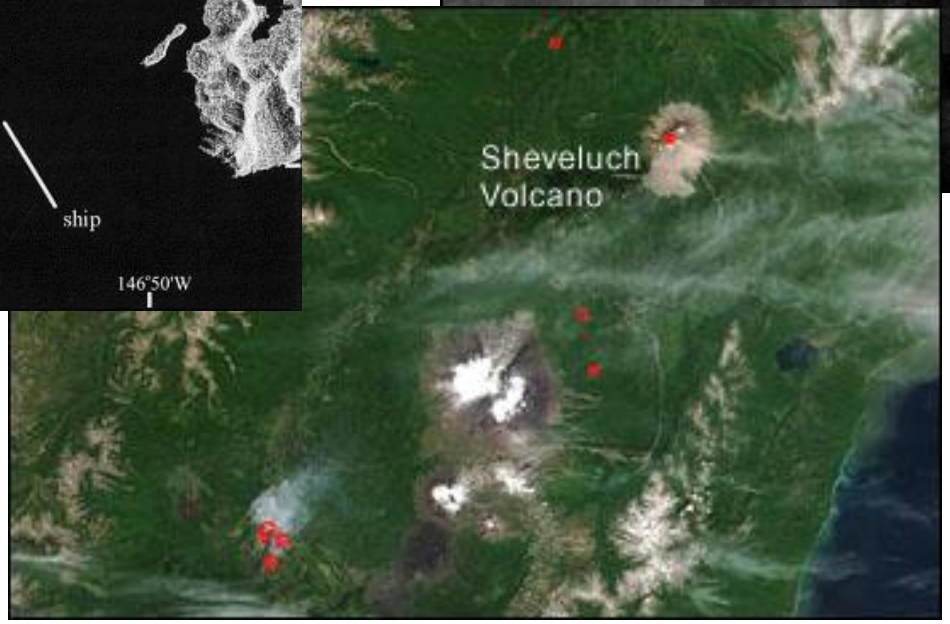
PENNSTATE

Other Missions and Sample Imagery



Icebergs





Forest Fires

PENNSTATE

Advantages of ImageScan

- Automated scanning to minimize manual review of all data and identify “high interest” data
- Automated target detection
- Automated reporting
- Automatically builds a historical database for trend analysis

PENNSTATE

Proposed Methodology

- Phase I – Design and Setup
- Phase II – Prototype Development
- Phase III – Automation Development
- Phase IV – Output Design & Implementation

Phase I – Design and Setup

- Choose a mission
  - Choose a mission with target(s) of interest that can be observed with bright on dark or dark on bright contrast
  - Identify a sensor that will accomplish the mission
- Identify a data source with several raster images
  - Choose several data sets with targets of interest
    - Easy, high contrast, little or no issues or artifacts
    - Medium, high contrast targets, various sizes to test spatial size filter, different geographic area(s)
    - Hard, complex area, varying sizes and intensities
- Explore and choose user inputs and format
  - Inputs: Date time, Geographic area, Intensity range, Other
  - Formats: csv/text, shape & attributes
  - Input methods: GUI, text file, off-line or part of the program
- Identify initial software used
  - ArcGIS, Other?

PENNSTATE

Phase II – Prototype Development

- Develop a working prototype - Manual design and test cycle
  - Design and test target detection algorithm within ArcMap
  - Test target detection algorithm on initial/easy data sets
  - Identify issues with target detection algorithm
  - Refine algorithm to improve detection, speed, accuracy and retest

PENNSTATE

Phase III – Automation Development

- Automate prototype
  - Write code to run the prototype process from a command line script for one data set
  - Modify the code to work on a data set with any filename
  - Modify the code to work on a list of files

Phase IV – Output Design & Implementation

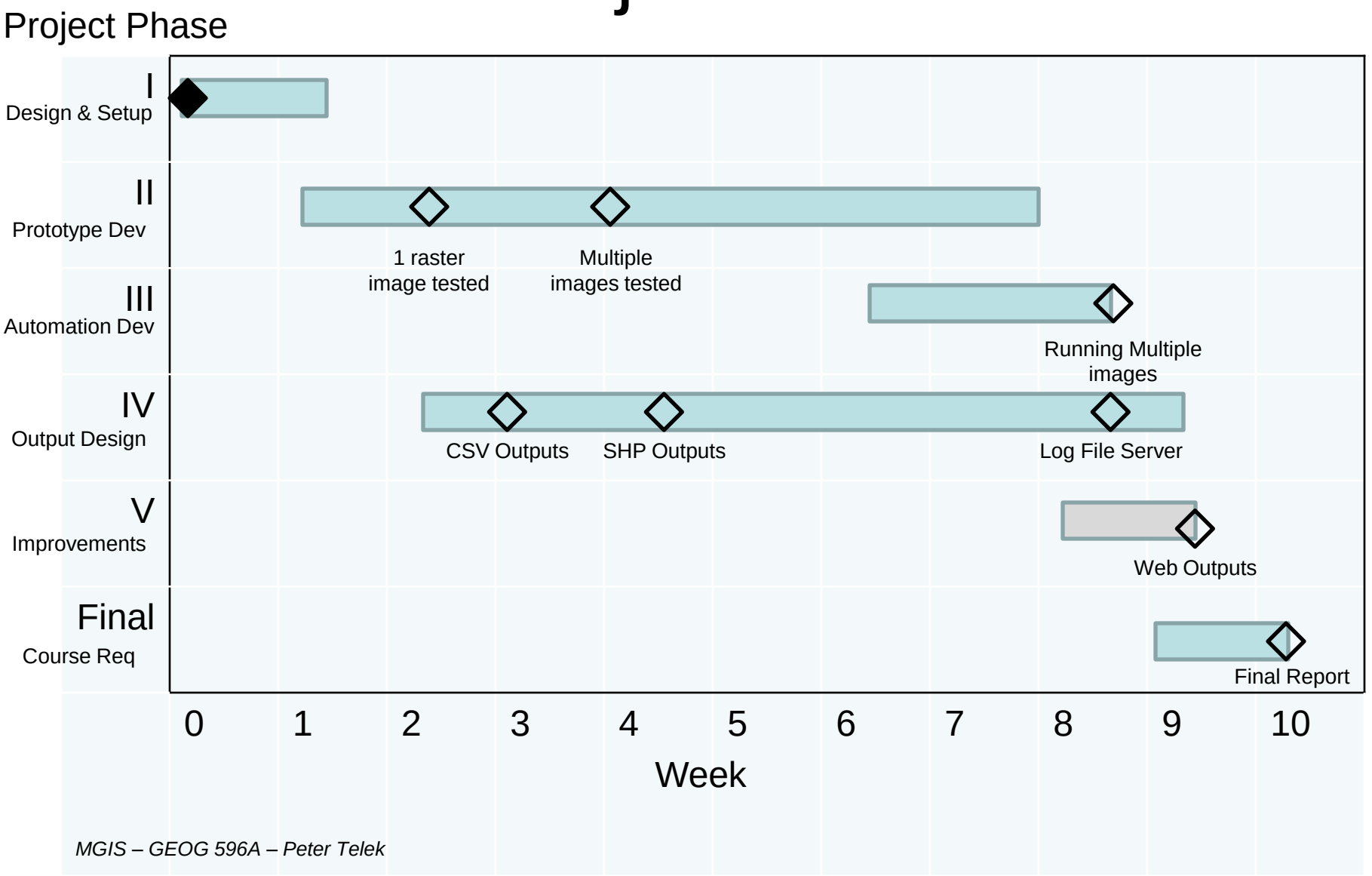
- Design the output
  - Identify any requirements
    - Information - target
    - Medium - files, intranet, internet, WMS/WFS, GoogleEarth, email, test msgs, smart phone
    - Information - Automated processing logs
  - Design and develop a sample output
    - Input data footprints
    - Target detections
    - Possible file formats (CSV, text, SHP, KML)
  - Design and create a summary report and metrics
    - If run on a large set of data (receive a hard drive of data with hundreds of data files)
    - If run daily (run routinely)
    - If run on images as they arrive (in an ingest folder)

Process Improvements *(as time permits)*

- Add additional user input parameters
- Add watch box alerts
- Change the detection algorithm
- Call an external algorithm detection program
- Add additional output

PENNSTATE

Project Timeline



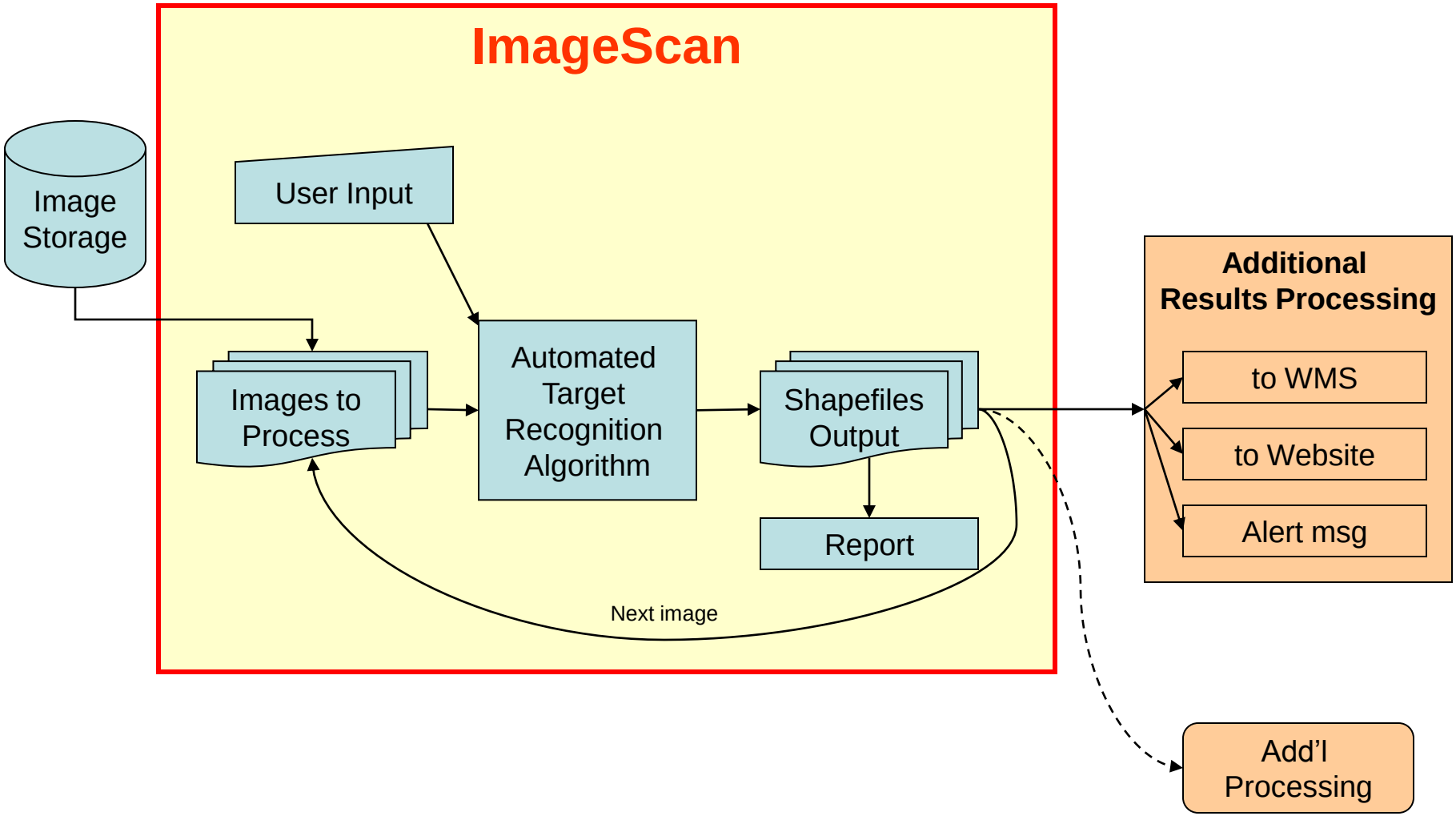
An example

- **Goal**: Detect and report locations of all probable hot spots with Thermal IR (MWIR) imagery

- Derived requirements:
  - Accomplish daily
  - Accomplish quickly with automated processing and minimal analyst input
  - Results will be in standard geospatial formats

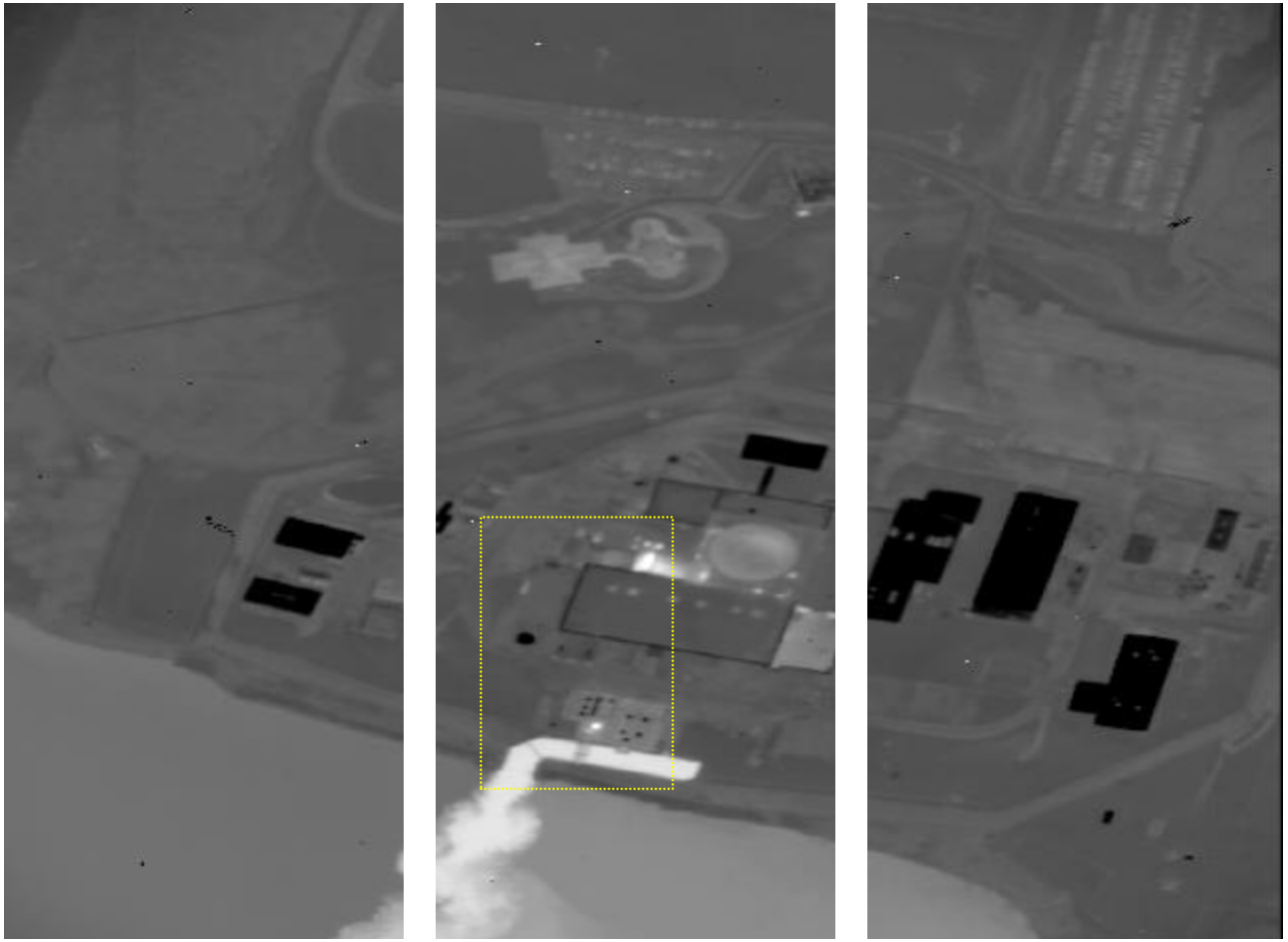
PENNSTATE

Anticipated Results - Process Overview



PENNSTATE

Sample Image



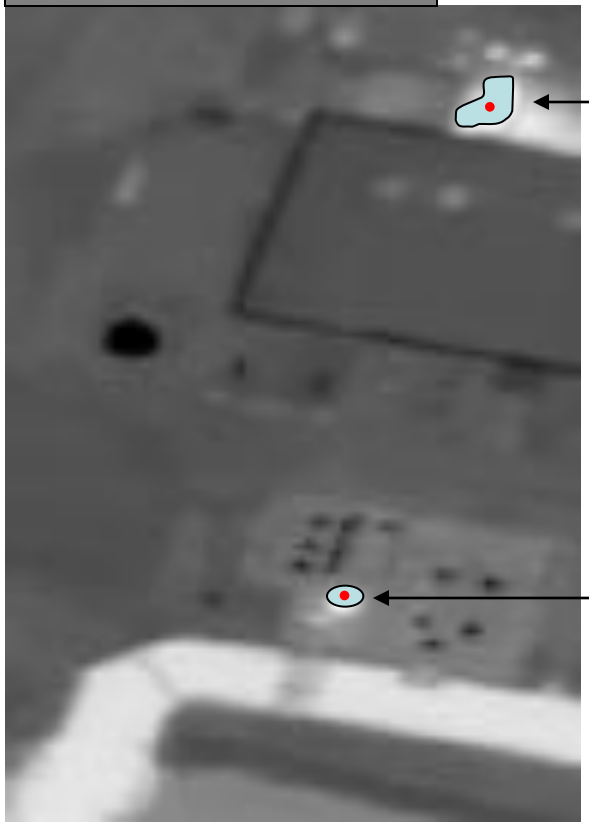
WASP Thermal MWIR Image of Ginna Power Plant

http://www.google.com/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=4&ved=0CDUQFjAD&url=http%3A%2F%2Flandsat.usgs.gov%2Fdocuments%2F6a_Schott_Overview_of_DIRS_Research%2520Activities.pdf&ei=N33JTqLfNa7y2gW6q4ngDw&usg=AFQjCNHKb35F1bXIjMoikF9xyUvBxEInCQ&sig2=vop2MSeAqS7z707GnHjOaw

PENNSTATE

Anticipated Results

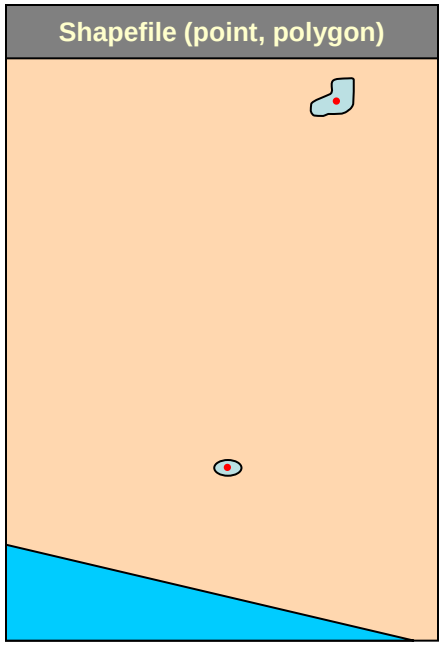
Image Close-up



Detection #: S001
ImageID: 20111116_041234.nitf*
Date: **2011-11-16***
Time: 04:12:34Z*
Lat: 43°16'43.25"N
Lon: 77°18'30.80"W
MaxValue: 1024*
Area: 12 sq m*
Sensor: WASP (MWIR)
Notes: Sample Automated Detection - Not verified
* *sample values*

Detection #: S002
ImageID: 20111116_041234.nitf*
Date: 2011-11-16*
Time: 04:12:34Z*
Lat: 43°16'43.25"N
Lon: 77°18'30.80"W
MaxValue: 1024*
Area: < 3 sq m*
…

Shapefile (point, polygon)



GoogleEarth KML

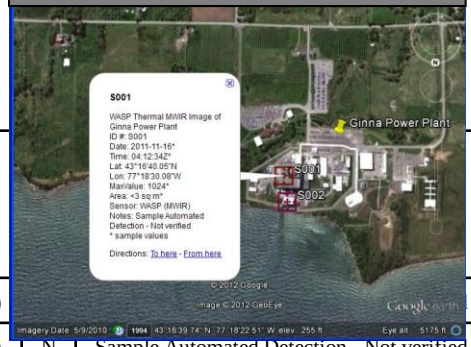


Table (CSV files)

ID	IMAGEID	DATE	TIME_Z	LAT	LON	MAXVALUE	AREA_SQ KM	SENSOR		
S001	20111116_041234.nitf	2011-11-16*	04:12:34Z*	43.27868056	-77.30855556	1024*	12.34	WASP (MWIR)		
S002	20111116_041234.nitf	2011-11-16*	04:12:34Z*	43.27868056	-77.30855556	1024*	2.2	WASP (MWIR)	N	Sample Automated Detection - Not verified

PENNSTATE

Summary

- Data volume is increasing
- Identifying smarter ways to process, store, transmit, and analyze large volumes of data is essential
- The proposed ImageScan process will be an ArcGIS based tool to complete an unsupervised processing and reporting chain to examine georectified raster imagery to automatically
  - Identify high interest data sets and
  - Report and database locations of potential targets of interest
- Shapefile and tabular results can be used for workload management, trend analysis, watch box reporting, and alert reporting

- Request approval to proceed with the proposed ImageScan project

PENNSTATE

Acronyms & Definitions

- AGI - Advanced Geospatial Intelligence
- ArcMap - ESRI Geospatial Analysis software
- ATR - Automatic Target Recognition
- BATC - Ball Aerospace & Technologies Corp.
- CBP - Customs and Border Patrol
- CSV - Comma Space Variable text file – common text file format for tables
- DHS - Department of Homeland Security
- DN - Digital Number
- ENVI - Commercial Remote Sensing Analysis software
- GEOINT - GEOspatial INTelligence
- HTML - Hyper text markup language - common file format for web pages
- IC - Intelligence Community
- L0/L1/L2/L3 - Level # processing (L0 is raw ◊ L3 is substantially more processed)
- NGA - National Geospatial Agency
- NITF - National Image Transfer Format
- NORTHCOM - Northern Command
- Opticks - Open Source Remote Sensing Analysis software
- shp - shape file (ESRI GIS standard) - – common GIS file format
- TIR - Thermal Infra Red
- USCG - United States Coast Guard
- USFS - United States Forest Service
- WFS - Web Feature Service
- WMS - Web Mapping Service

PENNSTATE

References

- Costlow, Terry. *"Geospatial intel evolution will continue at rapid pace"*. Created Oct 19, 2011, retrieved Nov 24, 2011, from http://www.defensesystems.com/Articles/2011/10/19/GEOINT-technology-development.aspx
- Kehler, General C. Robert. *"2011 GEOINT Symposium Speech".* Oct 19, 2011, retrieved Nov 24, 2011, from http://www.stratcom.mil/speeches/2011/75/2011_GEOINT_Symposium_Speech/
- Raytheon Company. *"Detecting the Undetectable: Hyperspectral Sensors Meet Critical Military and Civilian Needs".* Updated Aug 11, 2011, retrieved Nov 26, 2011, from http://www.raytheon.com/newsroom/technology/rtn11_auvsi/hyperspectral/
- Rosenberg, Zach. *"Raytheon's ACES HY enters production".* Jun 19, 2011, retrieved Nov 26, 2011, from http://www.flightglobal.com/news/articles/paris-raytheons-aces-hy-enters-production-358260/NASA LANCE. "*Frequently Asked Questions (FAQ) - LANCE*". Updated Nov 28, 2011, retrieved Nov 29, 2011, from http://lance.nasa.gov/imagery/rapid-response/frequently-asked-questions/#faq03
- FIRMS. "*Fire Information for Resource Management System (FIRMS) Background."* Retrieved Nov 29, 2011, from http://maps.geog.umd.edu/firms/readmore.htm
- Friedman, Karen S, et al. *"Routine Monitoring of Changes in the Columbia Glacier, Alaska, with Synthetic Aperture Radar.",* Remote Sensing of Environment, Volume 70, Issue 3, December 1999, Pages 257–264, retrieved Jun 21, 2012, from http://www.sciencedirect.com/science/article/pii/S0034425799000425

Questions ?

Email: put122@psu.edu