GLYDER: Global Cyclone Detection and Tracking from Remote Satellite Data

Ashit Talukder (JPL)

Email: Ashit.Talukder@jpl.nasa.gov

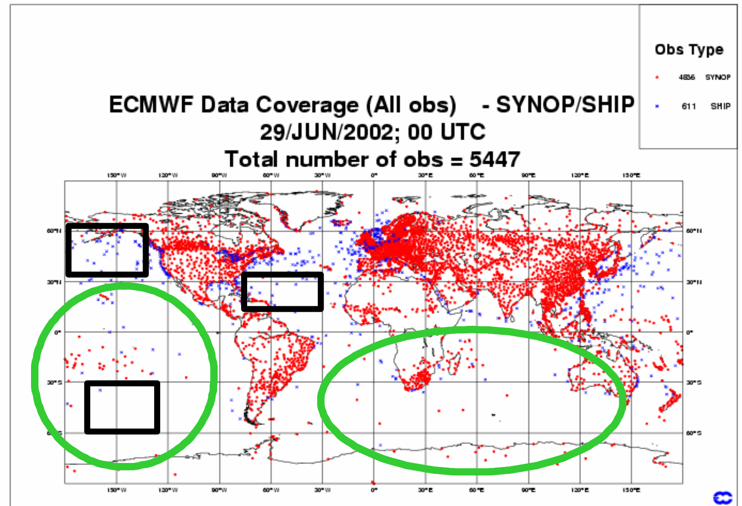
Ph: 818 354 1000

Anand Panangadan, Eric Rigor, Andrew Bingham, Tim Liu, Wendy Tang

Cyclone Detection: Current State of Art

Estimates of cyclone variability currently derived from analyses of surface level pressure (SLP) fields

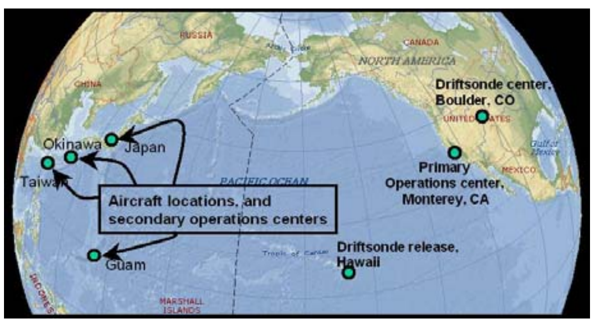
Model output based on in-situ inputs

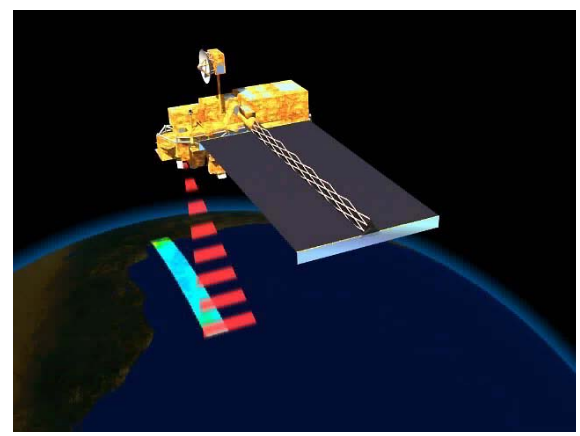


Map of daily pressure observations used by ECMWF Reanalysis forecast (6-29-02).

Extensive Field Campaigns to collect observations to track tropical cyclone in Western Pacific and South Eastern Pacific.

Observations still need to be manually analyzed

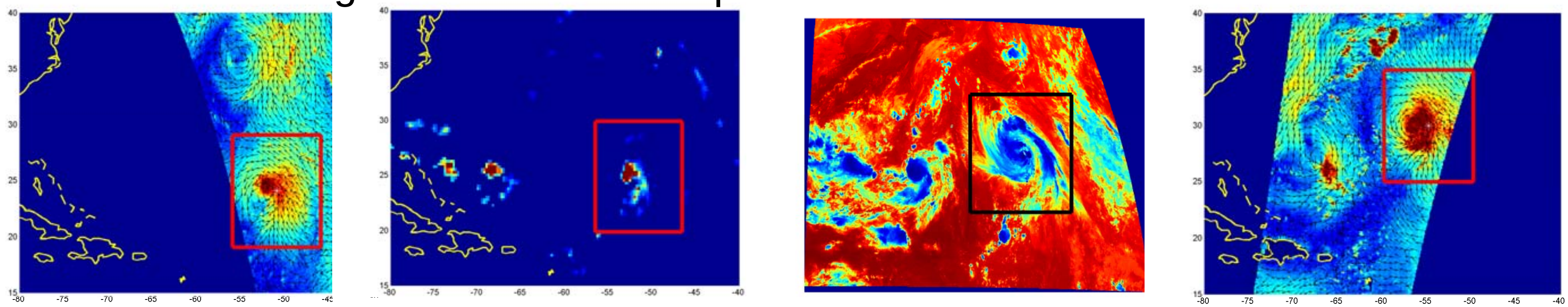




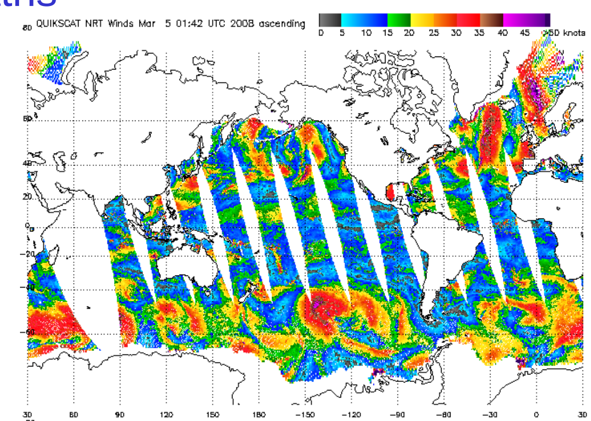
Satellite remote sensing provides global coverage potentially allowing detection of most/all global cyclones

Multi-Satellite Cyclone Detection Problem

- Several heterogeneous sources capture same event over time



- QuikSCAT – Wind measurements, Reliable detection, every 12 hrs
    - o Events occluded occasionally due to non-contiguous swaths

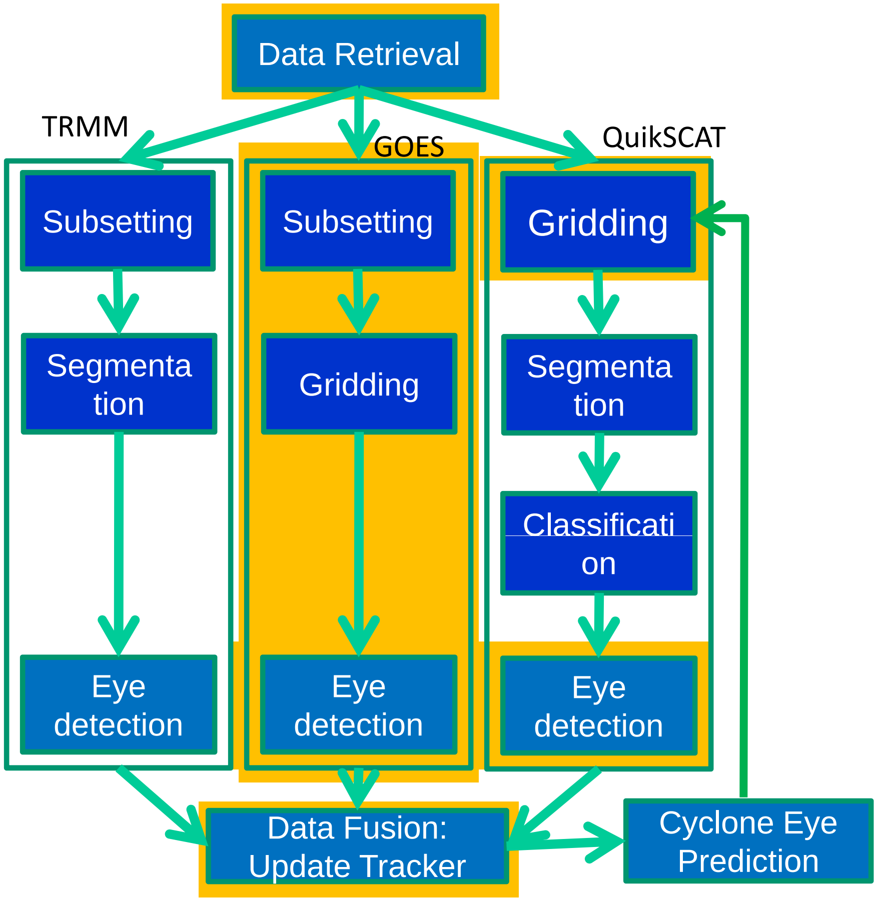


- TRMM – Precipitation, Weak detection, every 3 hrs
- GOES – IR, Cyclone structure, Reasonable detection, Every ½ - 2 hrs
    - o False alarms due to cloud cover

- Primary Solution
    - o Robust classification / detection from QuikSCAT windspeed
    - o Improve tracking temporal resolution and resolve occlusion using TRMM and GOES using predictive search and multisource classification

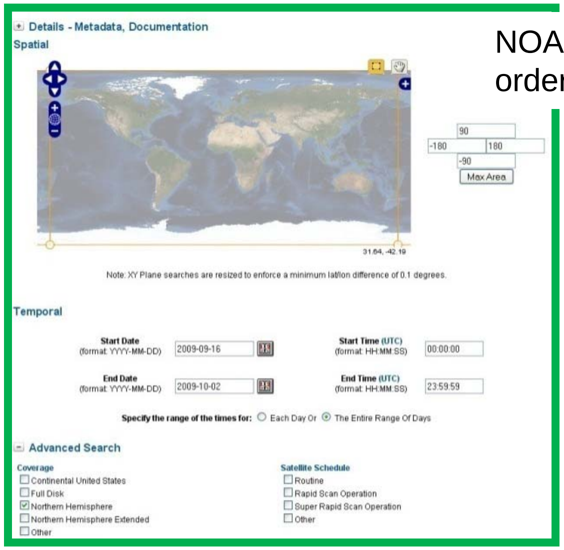
Improvements to the GLYDER system

- Incorporation of GOES data
  - Higher temporal resolution than earlier QuikSCAT and TRMM
- New techniques for eye detection
  - Log-spiral pattern matching for GOES images
  - Graph-based method for QuikSCAT
  - Generated classifier for Near Real-Time Data
- Predictive Multi-Cyclone tracker
  - Ability to track multiple cyclones
  - Kalman filter for constrained cyclone search
- GUI
  - Menu-driven, User friendly parameter and processing selection
- Error detection and recovery
- Automated Pre-processing steps
  - Downloading of data files from FTP servers
  - Automatic selection of NHC tracks (for error analysis)
  - Gridding of data
- Multisource Cyclone tracker
  - Share knowledge/information between disparate sources
  - Fundamentally new learning mechanism for multsource tracking
- Cyclone detection from Near real-time (NRT) data
  - New Processing to handle lower quality of NRT datasets
  - Assimilate RT track & prediction information from various sources (RSS feeds, website HTML pages, etc.)

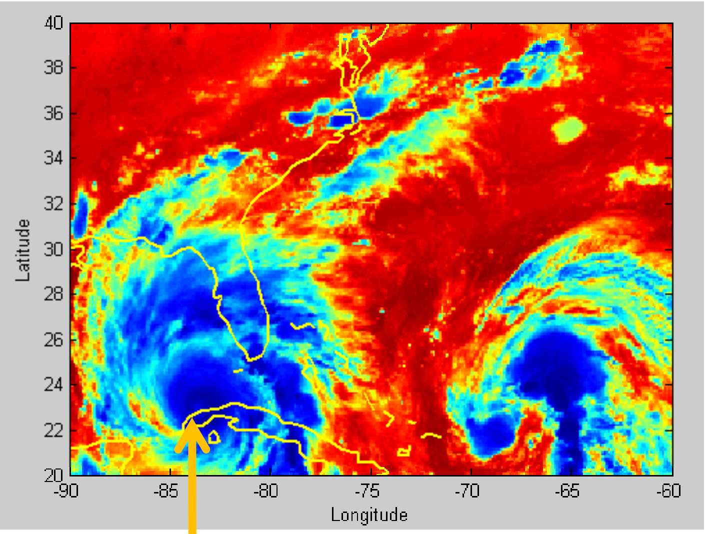


Geostationary Operational Environmental Satellite (GOES)

- Chief advantage: high temporal resolution
  - Images available at 30 minute intervals
  - Earlier version used
    - QuikSCAT: about 2 per day
    - TRMM 3B42: 3 hour intervals
- GOES-12 (located at 75W) for Atlantic coast
- GOES-11 (located at 135W) for Pacific coast
- We use Band 4: 10.2 µm – 11.2 µm Infrared
  - Observe storm clouds at night too
  - Spatial resolution: 4km at nadir
- Images downloaded from NOAA's CLASS library
- Challenges
  - Cyclones are not *defined* by cloud shapes
    - We use GOES imagery only after initial identification from QuikSCAT
  - Characteristic spiral pattern develops only when intensity increases
  - Large size of each image requires us to segment out the approximate cyclone location before calculating eye location



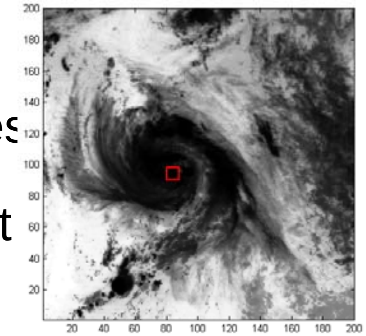
NOAA CLASS ordering webpage



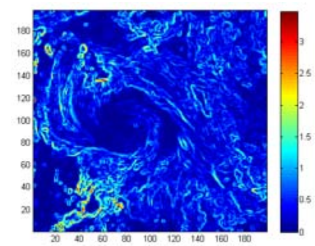
Log spiral cloud pattern

Eye location from GOES

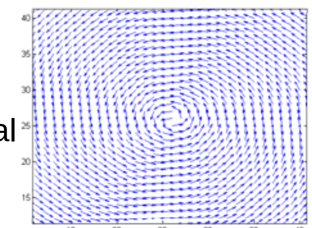
- Fit a log-spiral pattern over the image
- Pattern matching performed over the gradient images
- Score matrix: how well does pattern centered on that element match the image?
  - Element with highest score is returned as the eye location
  - Also implemented variant that compared the scores in every peak's neighborhood
    - Reduced error in cases where two spiral patterns were visible in search region
- Detected eye error decreases with time
  - As a hurricane intensifies over time, the spiral shape of the cloud bands becomes more prominent.



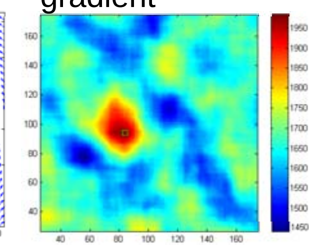
GOES image



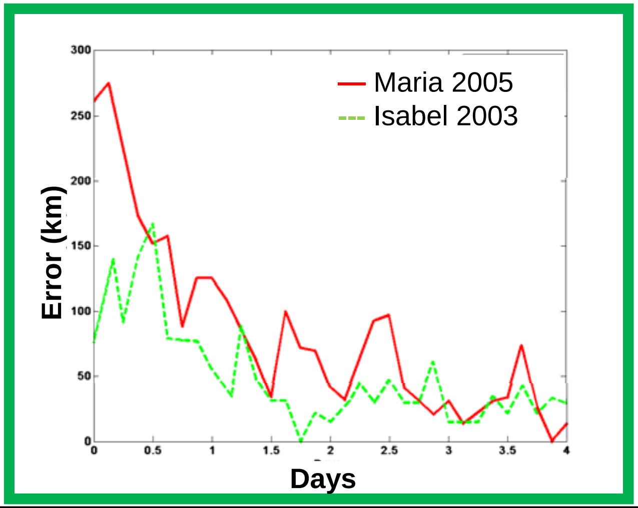
Magnitude of GOES image gradient



Log spiral pattern gradient



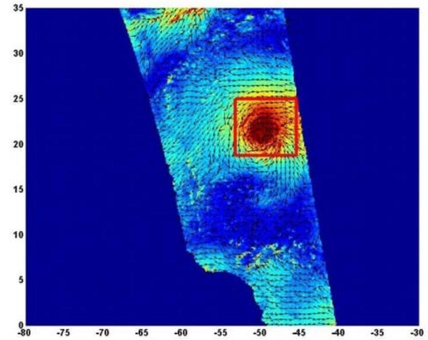
Pattern matching score matrix



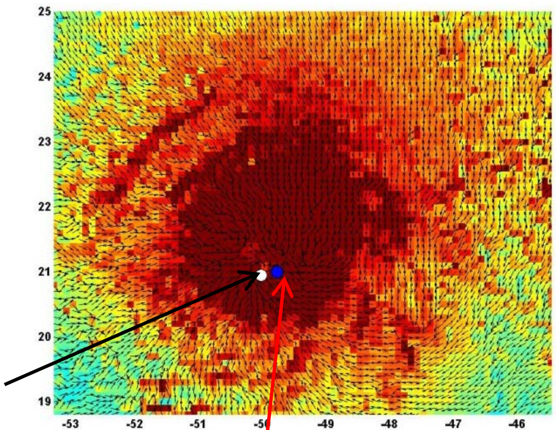
Eye location from QuikSCAT: Graph-based Method

Intuition: True eye should have largest number of wind vector pixels “leading” to it

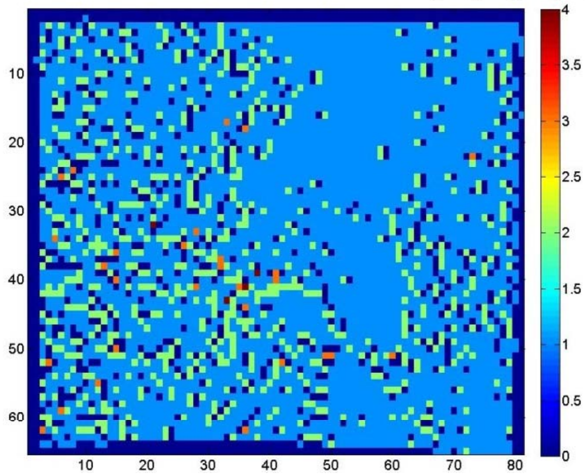
1 Segmentation and classification to locate cyclone region



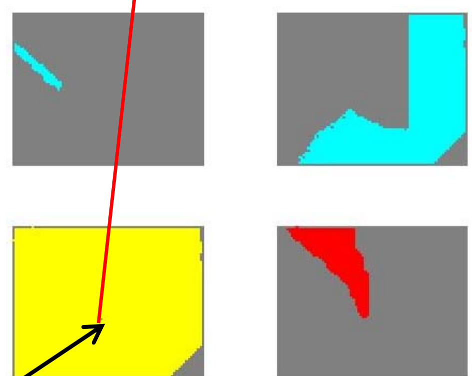
2 Compute normal vector to the directional vector for each pixel



3 Compute number of nearest neighbors pointing to each pixel



4 Construct spanning trees for pixels having the highest number of pointing neighbors

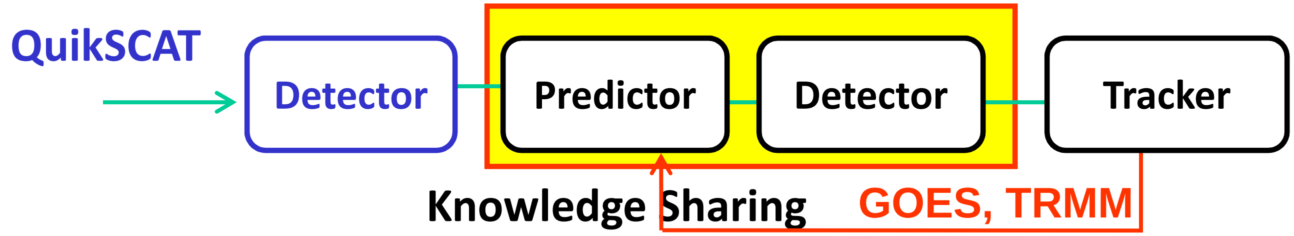


5 The pixel that grows the largest spanning tree is the cyclone eye

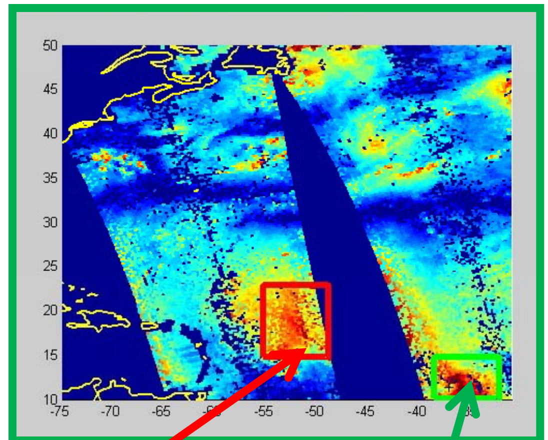
2X improvement in accuracy over earlier *centroid* method:

Hurricane	#images	Error: GB (km)	Error: Centroid (km)
Isabel 2003	21	87	175
Maria 2005	14	90	199

Predictive Cyclone Tracker



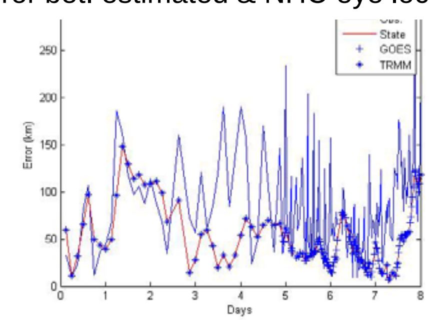
- The tracker provides **future cyclone location *estimates*** as new satellite data becomes available
  - This constrains the search region for the cyclone eye detection algorithms
  - A smaller search region
    - reduces the incidence of false positives
    - reduces the computational processing time



- Kalman Filter Model-based tracking
  - State model: how does cyclone eye move over time?
  - Observation model: how are observations (results of eye detection algorithms) related to true eye location?

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \\ v_{x,t+1} \\ v_{y,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ v_{x,t} \\ v_{y,t} \end{bmatrix} + \boldsymbol{w}$$

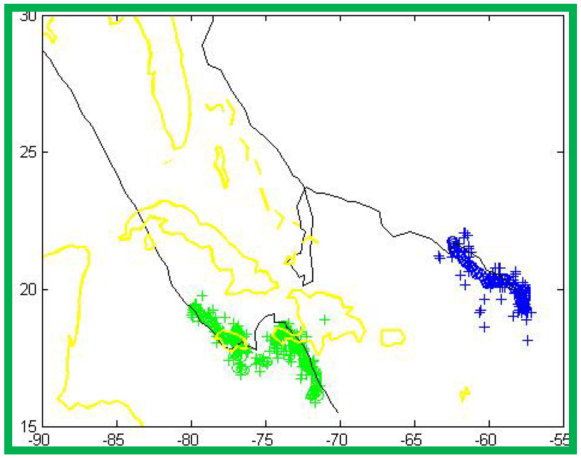
Error bet. estimated & NHC eye locations



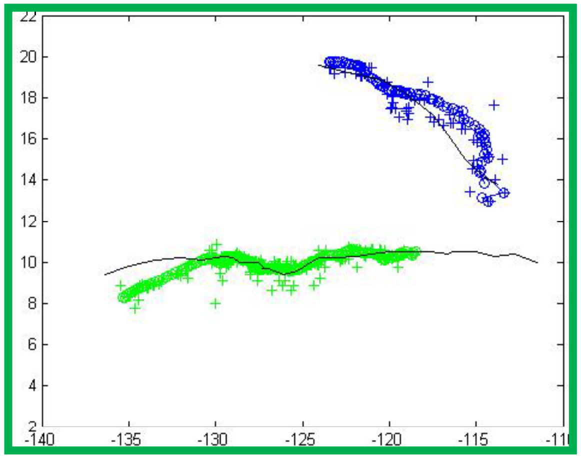
- Filter *state* estimates are smoothed versions of the observations (eye detections)
  - State estimates reduce the effect of random errors

Tracking Multiple Cyclones

- Earlier: the system assumed that only exactly one hurricane is represented in the images
- Current tracker can track multiple cyclones simultaneously
  - Separate filter is instantiated for each cyclone
- New challenge: *Correspondence problem* has to be solved
  - How to assign multiple observations to multiple tracked cyclones?
  - Detect new cyclone
    - distinguish from observation error
  - Detect when a cyclone has ended
    - Distinguish from temporary failure in observations
- Solutions: exploit characteristics of each data set
  - Instantiate new cyclones only from QuikSCAT
  - Use TRMM and GOES to propagate instantiated trackers
  - New parameters
    - Max. allowable distance between observations and state
    - Max. number of images without observing a tracked cyclone



Hurricanes Gustav, Hanna 2008



Hurricanes Carlos, Dolores 2009

Graphical User Interface (GUI)

- Enable all input data and system parameters to be specified easily
- Allow plotting and processing to proceed independently
- Modularized the code: separate the major steps (initialization, processing, and plotting)
- Automatic download of data files from relevant FTP servers
  - QuikSCAT and TRMM
  - GOES data is downloaded using NOAA's CLASS

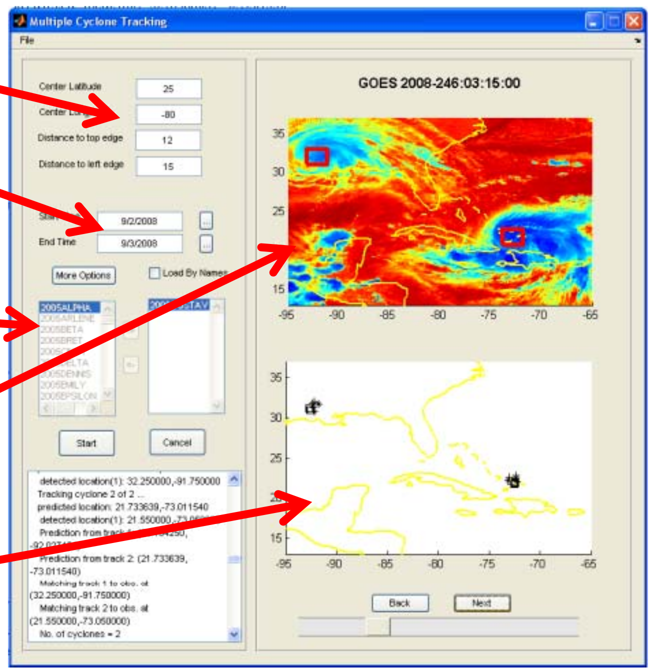
Search region

Start and end times

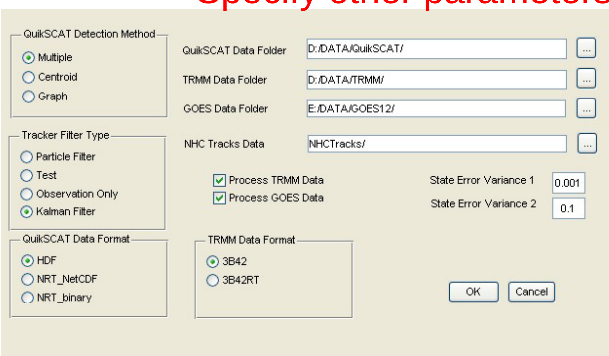
Select NHC tracks (ground truth)

Current image with detected eyes

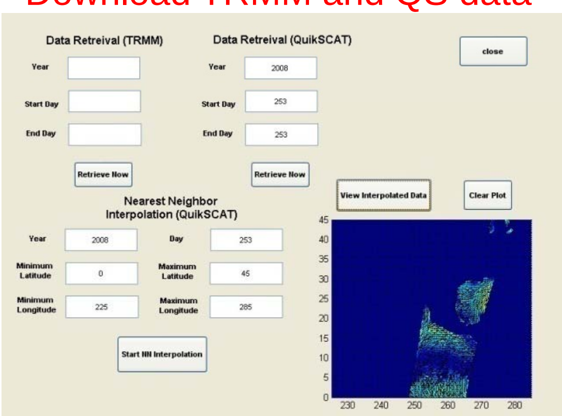
Estimated hurricane path



Specify other parameters



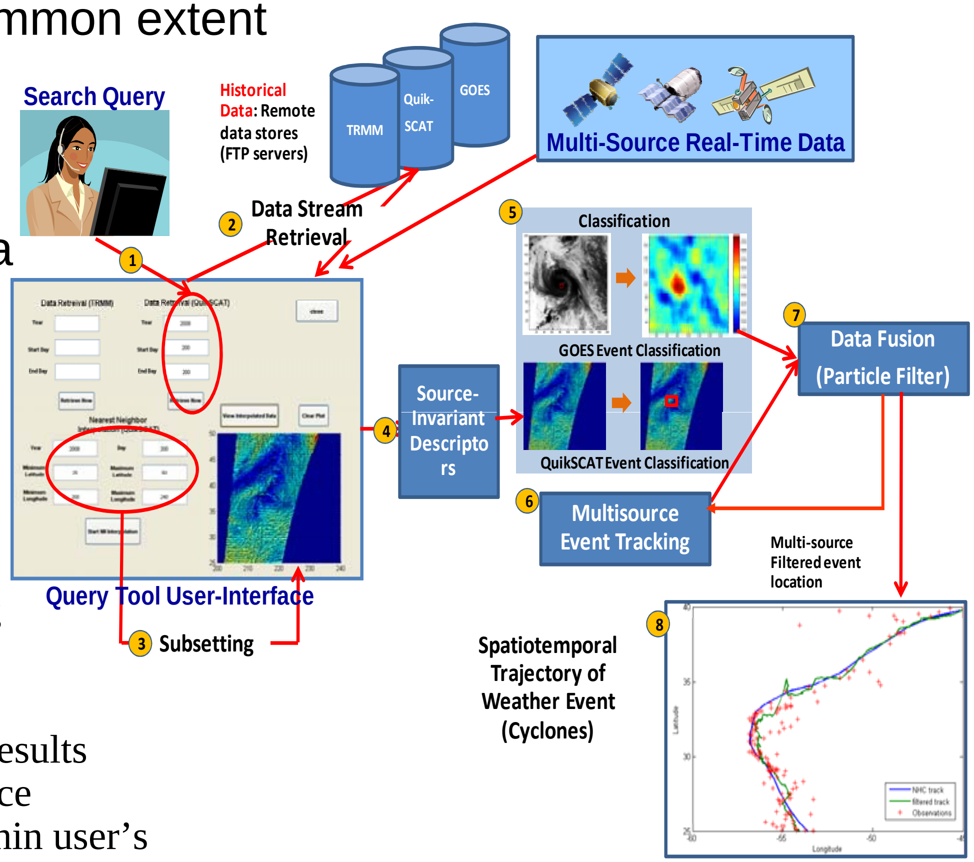
Download TRMM and QS data



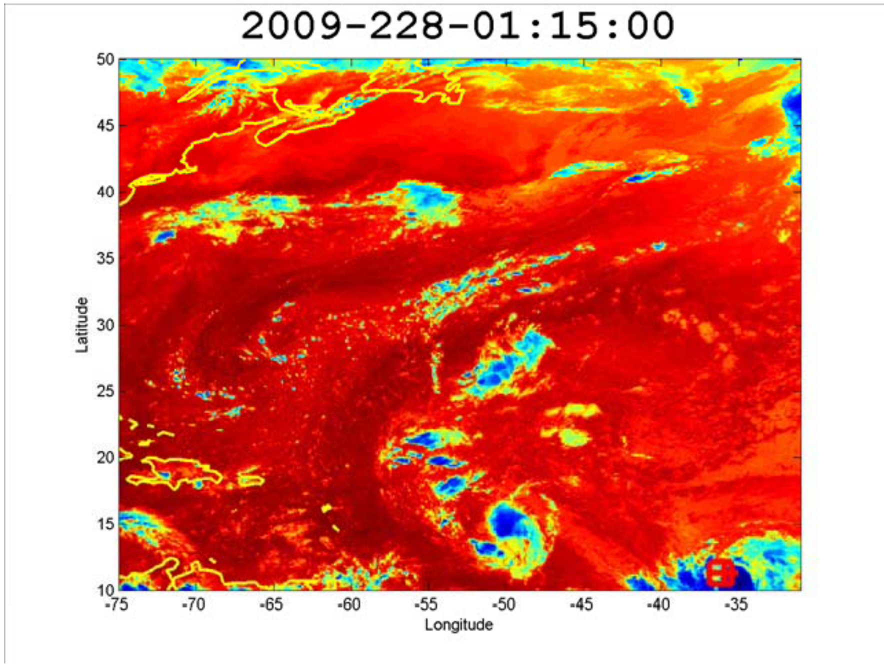
Other improvements

To improve usability and running speed

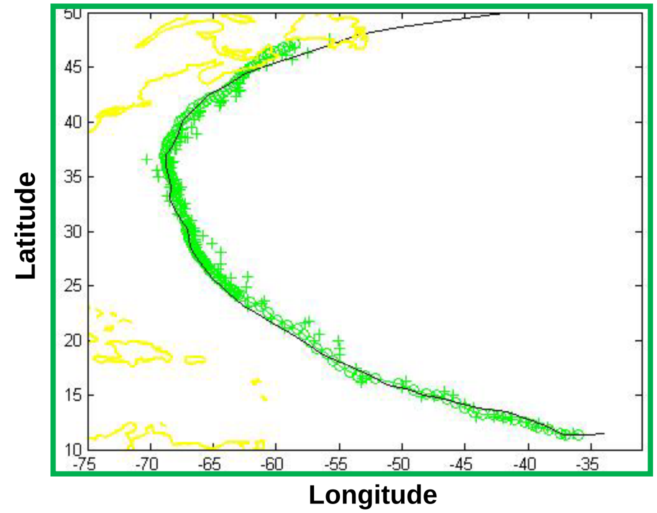
- Gridding and segmenting of all data types (GOES, TRMM, QuikSCAT) to a common extent
  - Enable post-processing of data
  - E.g..: feature-feature tracking
- Increased accuracy of GOES eye detection allows smaller search area
  - Increases running speed since gridding is slowest step (GOES, QuikSCAT)
- Error detection and recovery
  - Check for feasibility of input parameters
  - Recover from unexpected run-time errors
    - Out-of-memory, missing data files
    - Skip current image and continue processing
- Automatic selection of NHC tracks
  - NHC tracks are used only for analysis of results
  - Downloaded all NHC best tracks in advance
  - Automatically select only those tracks within user's space and time input range to calculate errors



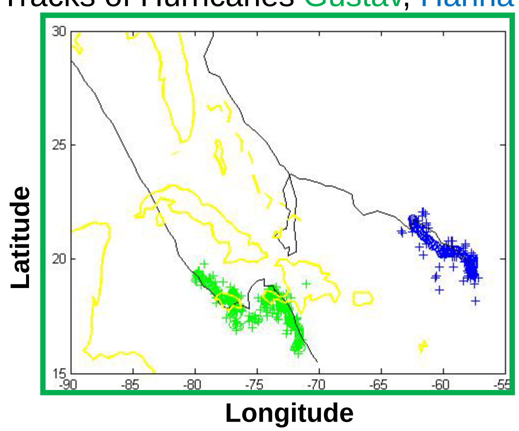
Results on 2009 Cyclones

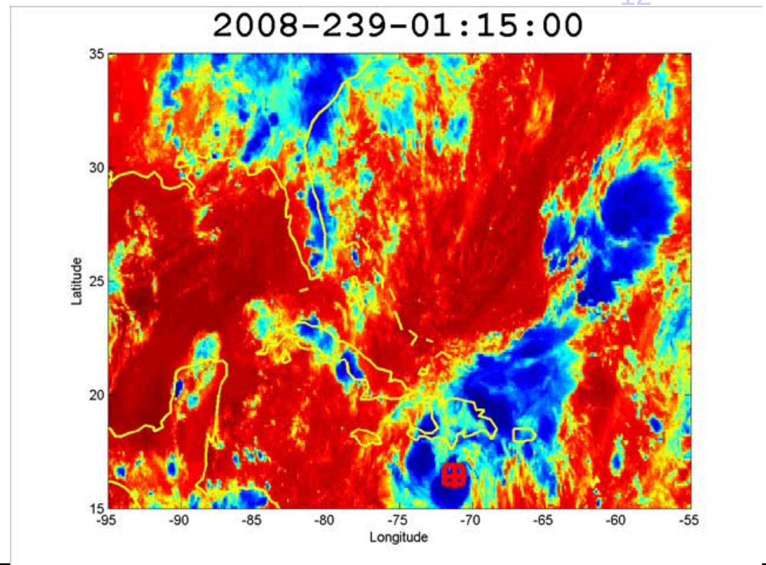


Detected and NHC track of Hurricane Bill 2009



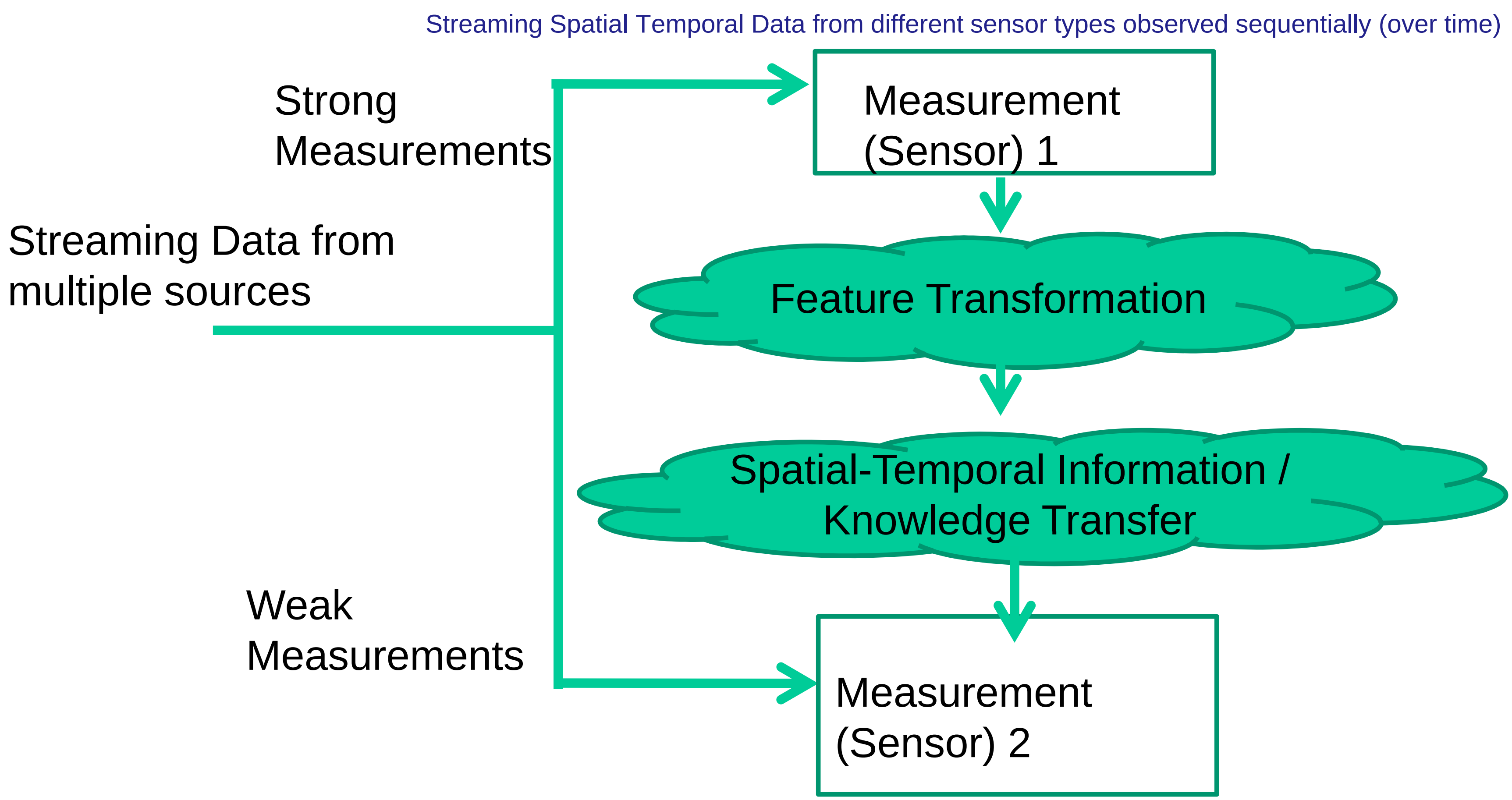
Tracks of Hurricanes Gustav, Hanna 2008





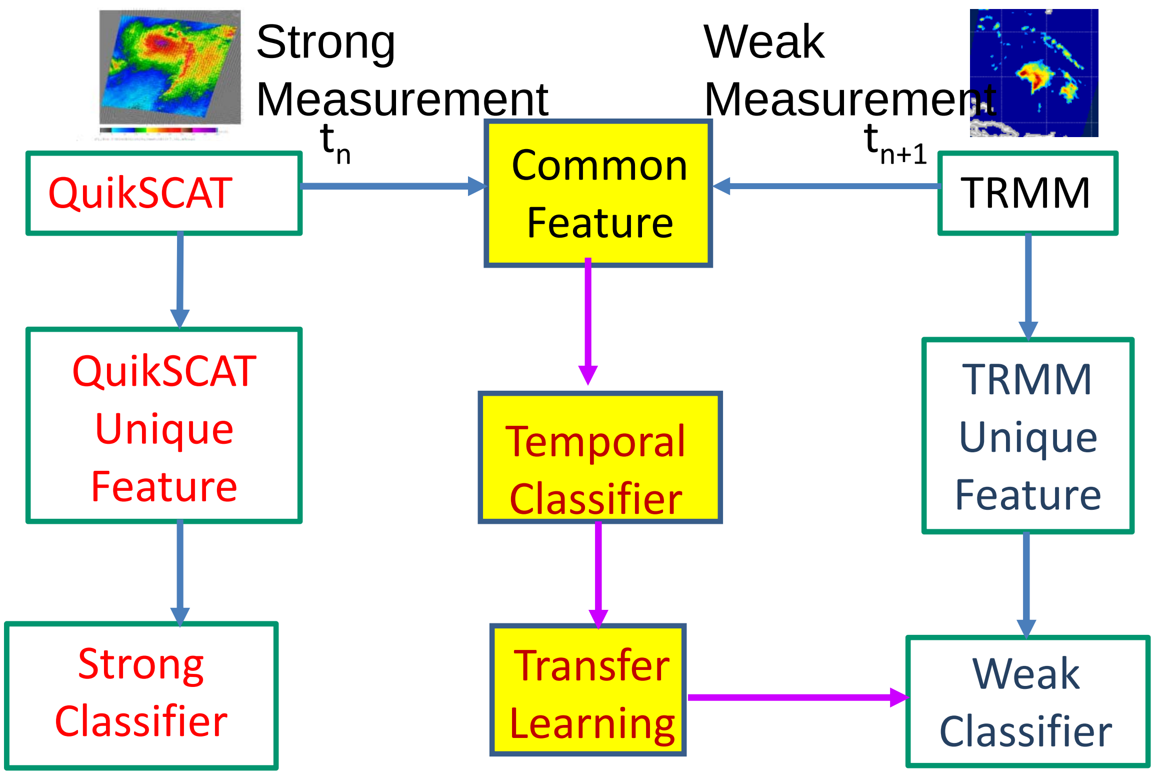
Multisource Knowledge Transfer Overview

- Predictive tracker provides constrained search region for next observation
- Should leverage sequential observations from multiple sources for more robust tracking



MultiSource Classification Solution

Transform data into different component (feature) spaces for multisource classification over time and space



Problem Definition: Generative Model

- **Goal is to detect and classify the event over a length of time t to $t+\Delta t$ in space *s* which is represented by the set $I(s_{t,t+\Delta t})$.**

- *Many prior machine learning and transfer learning solutions* for learning and classification from multiple sources (say *X, Y*) *assume that simultaneous measurements from all sensors (measurement sources) are available at all times*, i.e. pairs of observations exist at every instance $\{X(s_{t_o}),Y(s_{t_o})\},\{X(s_{t_1}),Y(s_{t_1})\}$ etc.
  - Most practical scenarios involve un-co-registered measurements at different times *t* and different locations *s*.

- Observed sequence has alternate measurement observations from the two sensors as follows $X(s_{t_o}),Y(s_{t_1}),X(s_{t_2}),Y(s_{t3}),\ldots\ldots\ldots X(s_{t_N}),Y(s_{t_{N+1}})$
  - Existing multisource learning techniques (canonical corr, discriminative solns) will not work

- **Model multisource learning and classification as a generative process where observations are comprised of two components:**
  - **(a) components $C(s_t)$ that are common to all (both) data sources, and**
  - **(b) components $U(s_t)$ unique to each data stream**

Problem Definition: Multi-Source Knowledge Transfer

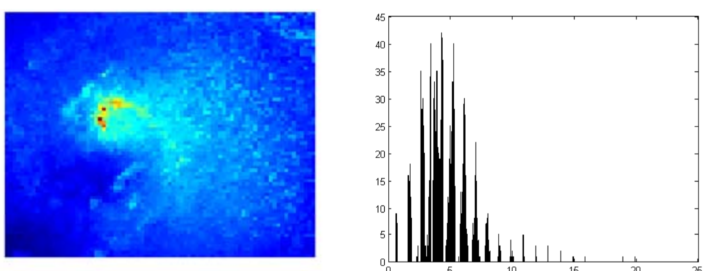
- **Use an invertible data representation for multi-source measurements**
  - the components of each source measurement are orthogonal to each other
  - the transformation is linear in nature. Where

$$X(s_{t_o}) = C(s_{t_o}) + U_X(s_{t_o}); \; Y(s_{t_1}) = C(s_{t_1}) + U_Y(s_{t_1});$$
$$C(s_{t_o}) \perp U_X(s_{t_o})$$

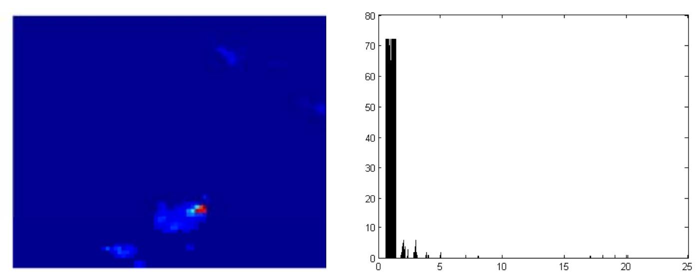
- ***Multi-phase classification using common & unique features from multiple-sources***

- ***Transfer knowledge between strong source measurements at*** t_0 ***and weak measurements at*** t_1 ***using common feature space*** $C_X(s_{t_o}), C_Y(s_{t_1})$

- Multi-source knowledge transfer involves a pairwise classification of successive observations from sources X and Y $L\{(C_X(s_{t_o}), C_Y(s_{t_1})\}$



Quikscat (at time t0)



TRMM (at time t1)

Problem Definition: Multi-Source Classification

- **Unique Component Classification in multi-source measurements**
  - Compute unique components $U_X(s_{t_0})$, $U_Y(s_t)$ for each source *X*, *Y*
  - Use source-specific classifiers $W\{\}$ to label each data stream using the unique components $U(s_t)$

- **Each data source differs in discriminative capabilities for event classification**
  - Some data sources (say X) have strong discriminative power with a strong classifier using the unique features $W\{U_X(s_{t_0})\}$
  - Other sources (say Y) have limited discriminative capabilities and weak classifiers $W\{U_Y(s_{t_1})\}$
  - Use common feature classifier $L\{(C_X(s_{t_o}), C_Y(s_{t_1})\}$ to transfer knowledge to weak classifier $W\{U_Y(s_{t_1})\}$.

- **Generalizes easily to multiple data sources (greater than 2)**

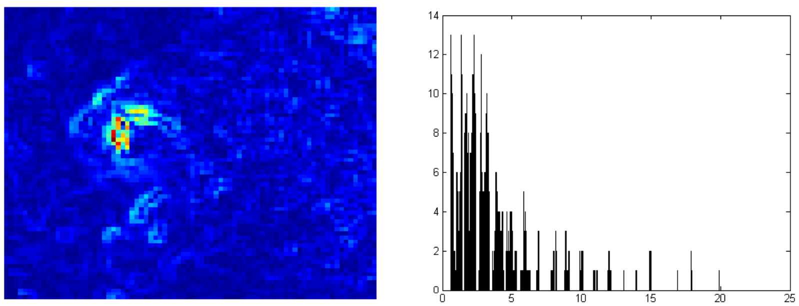
Problem Definition: Feature Space Computation

- **Use mutual information metric to estimate the transformations for common feature and unique feature spaces.**
  - Mutual information (joint-entropy) between the X and Y is maximized
    - Common feature is the transform that maximizes the common or shared information between X and Y.

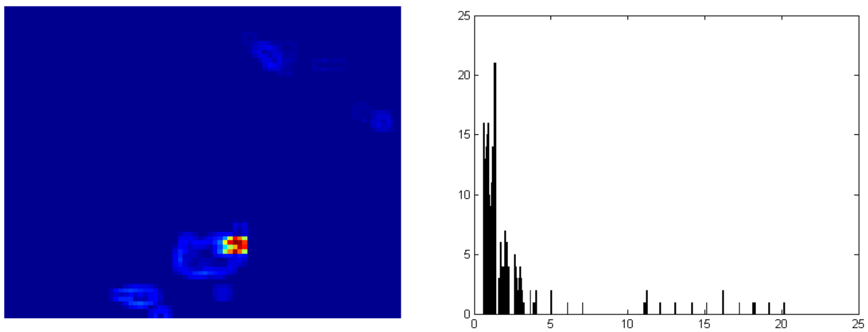
$$M(\mathrm{F}(X),\mathrm{F}(Y)) = H(\mathrm{F}(X)) + H(\mathrm{F}(Y)) - H(\mathrm{F}(X),\mathrm{F}(Y))$$

- **Transform can be linear or nonlinear**
  - Currently only linear transforms considered

- **Several linear and orthogonal transforms possible**
  - Wavelets , Fourier transforms (high-pass, low-pass, band-pass)
  - Principal component analysis



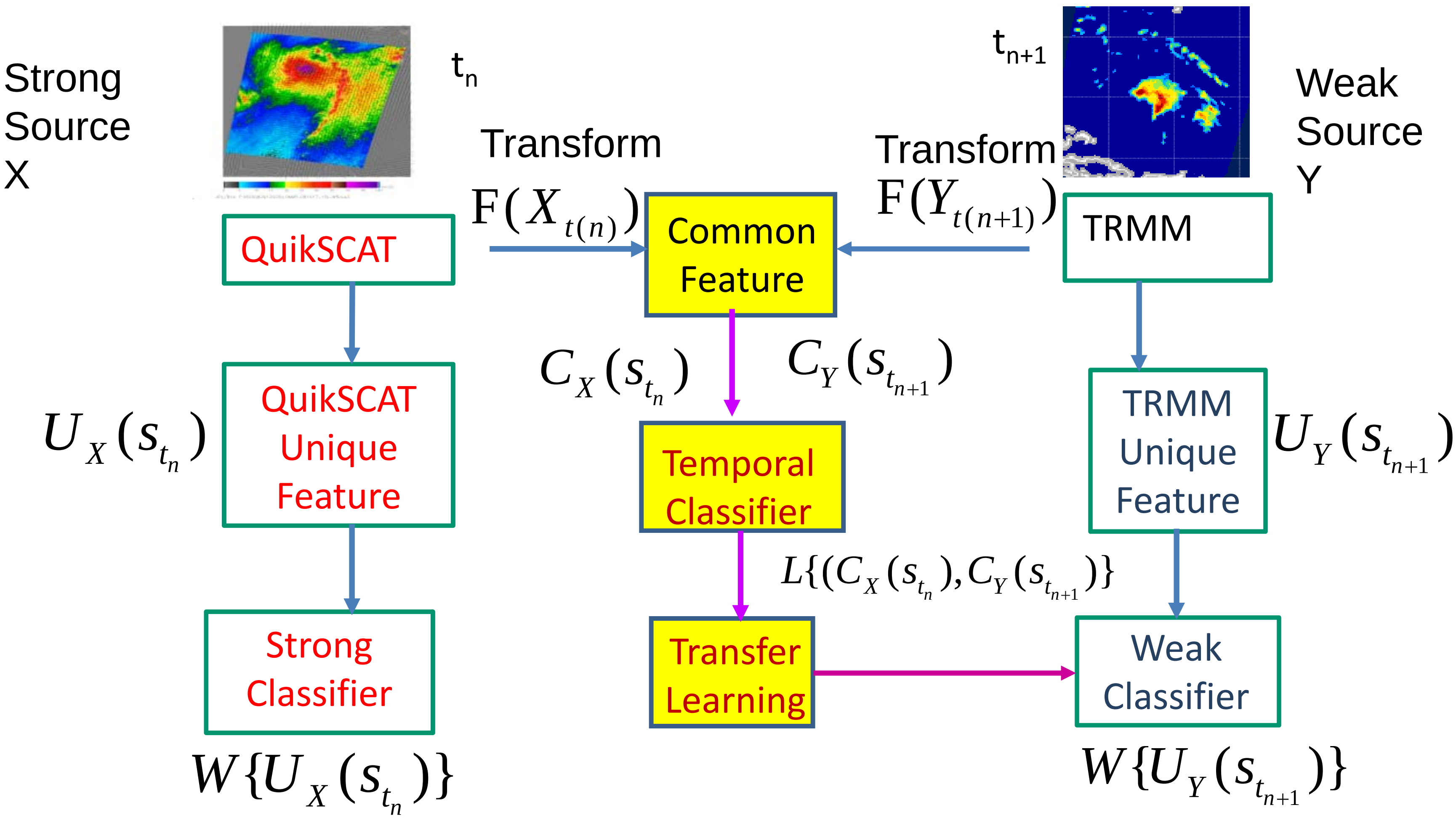
Quikscat common feature transformation (time t0)



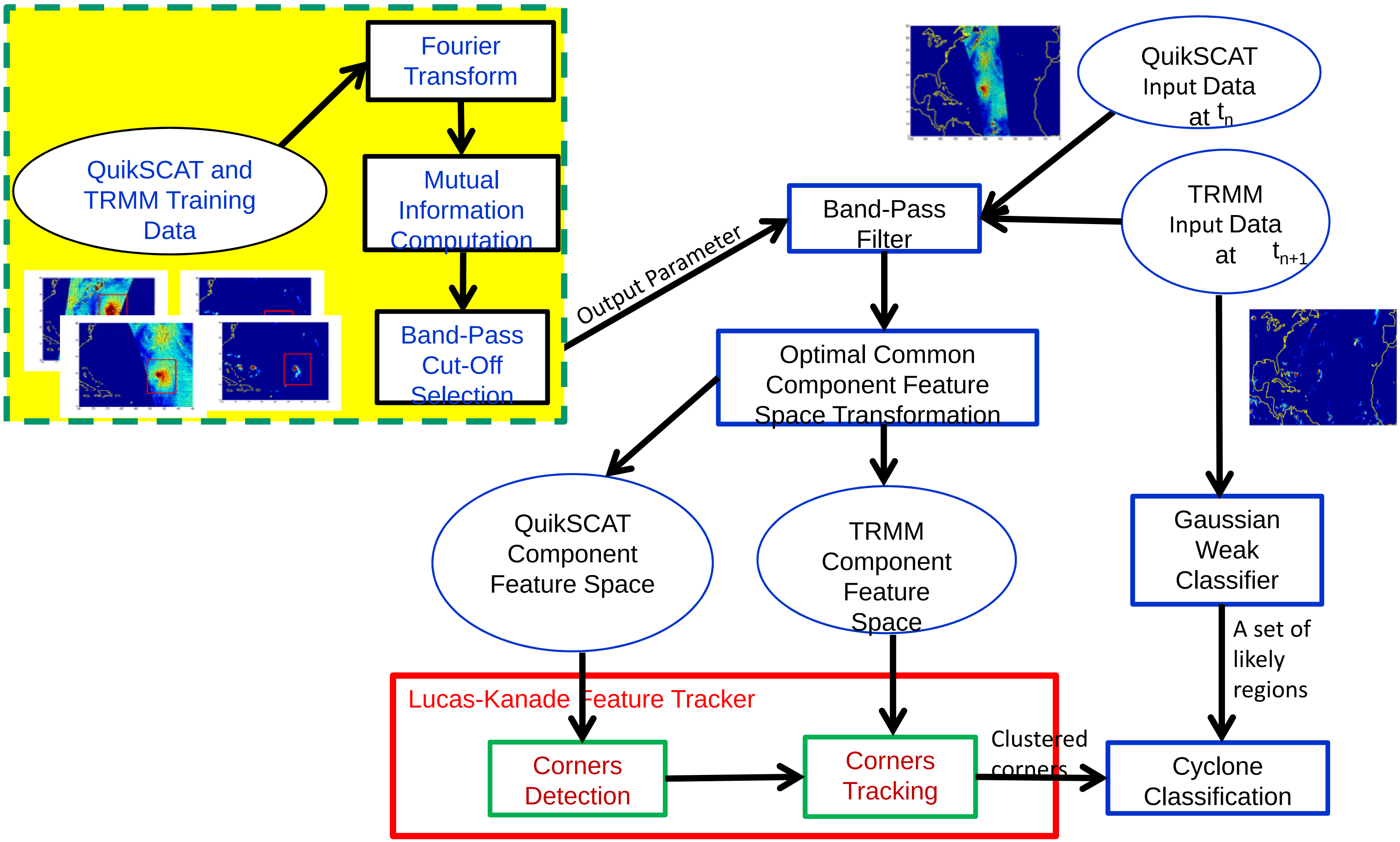
TRMM common feature transformation (t1)

MultiSource Classification Primary Steps

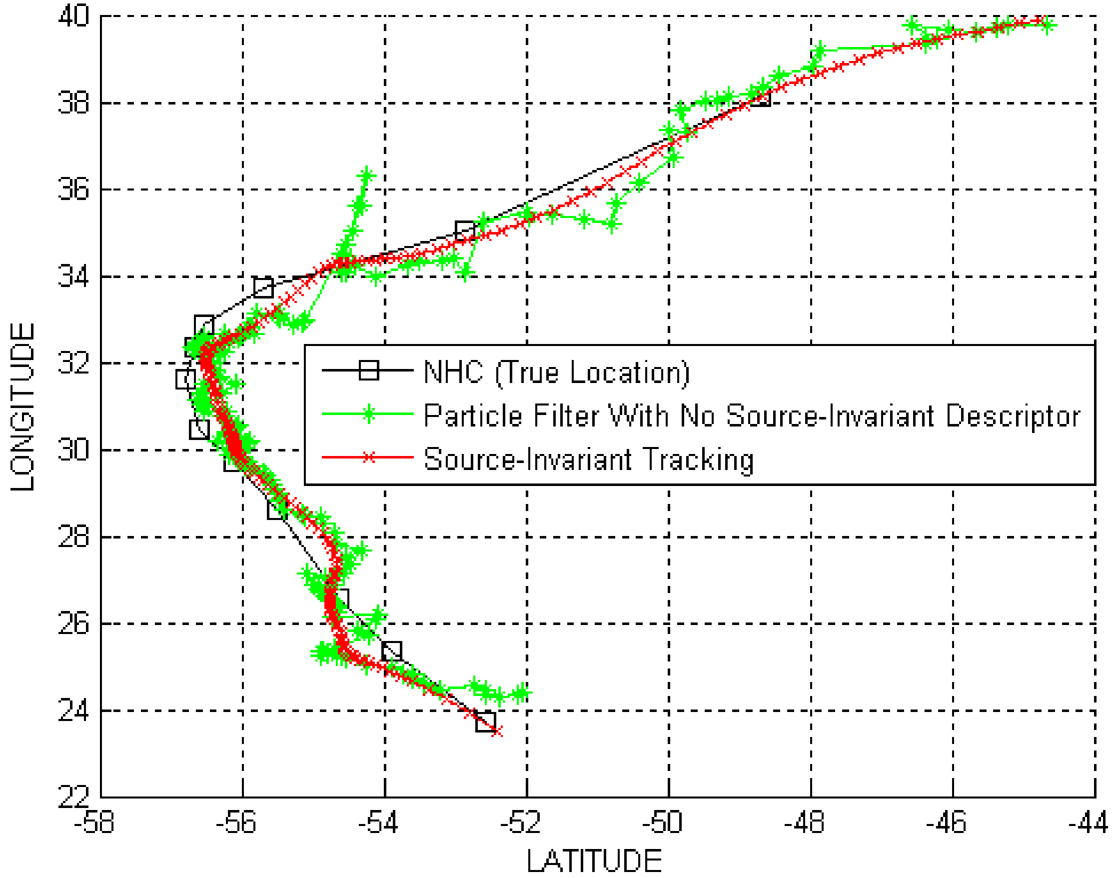
Transform data into different component (feature) spaces for multisource classification over time and space



MultiSource Classification Algorithm



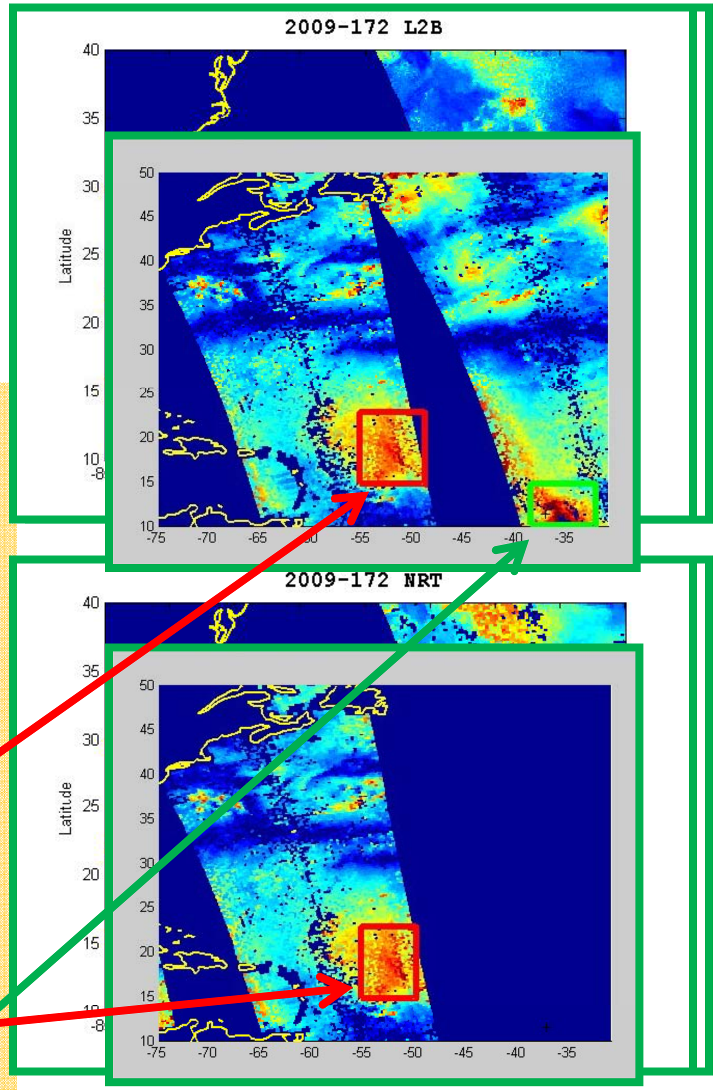
MultiSource Classification Results



Near-Real Time Cyclone Detection: Challenges

- 3B42RT provided by TRMM Science Data and Information System and GSFC Lab. for Atmospheres
- Combination of High Quality (3-hourly, 0.25°x0.25°) and calibrated Variable Rainrate
- Relatively minor differences in precipitation rates between 3B42 and 3B42RT
  - Same eye detection algorithms for both 3B42 and 3B42RT data

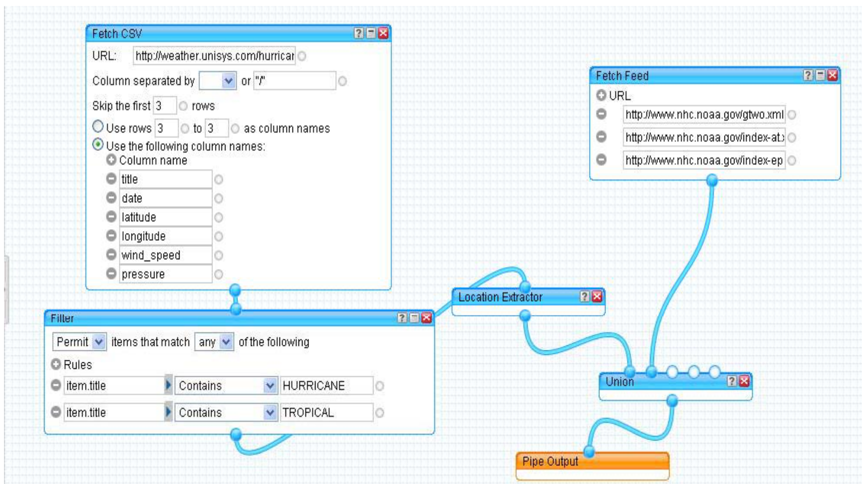
- Earlier version of GLYDER used QuikSCAT L2B data
  - High quality since initialization of the ambiguity removal algorithm is done using actual analysis fields in data
- SeaWinds Real-Time MGDR data product
  - Available within 1-3 hours of the satellite observation
- Significant differences between L2B and NRT images
  - Earlier classification algorithm shows higher number of both False positives and False negatives
- New classifier for NRT QuikSCAT
  - Created separate training datasets for the SVM-based hurricane classifier
  - First, identify potential segments containing a storm
    - Based solely on wind velocity magnitude
  - Positive examples: wind vector segments containing a hurricane eye (NHC)
  - Negative examples: segments with no eye

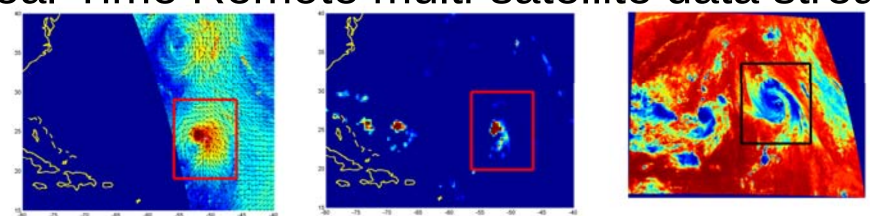


Near Real-Time Cyclone Tracking

- “Mashup” near real-time cyclone location information from numerous sites
  - NHC, UniSys and others report various cyclones in near real-time
  - Predictions of future locations also provided upto 48-72 hours in advance
  - Yahoo Pipes enables mashup of online information from HTML, XML,RSS

Real-Time Remote multi-satellite data stream





Single source location

Multi-source Tracking Estimate

Manual track prediction (multiple Online sites)

Assimilated cyclone location report

Future Work

- Improve detection algorithms from QS, TRMM, GOES
- Incorporate geostationary satellite data from other missions:
  - Meteosat, Kalpana; Track West Pacific and Indian ocean cyclones
- Improve, integrate multisource classification alg. from multi-satellite streams
- Improve, integrate near-real-time processing (include NHC data mashups)
- Release GLYDER: Host GLYDER system on a public-accessible website
  - GIOVANNI
  - Initial release will be as a self-contained executable to be run on the user's computer
- Publications
  - "Classification from Disparate Multiple Streaming Data Sources", Workshop on Learning from Multiple Source, held at Neural Information Processing Systems (NIPS-2008), December 2008.
  - "Automated Historical and Real-Time Cyclone Discovery With Multimodal Remote Satellite Measurements", Eos Trans. AGU, 89(53), Fall Meet. Suppl., Abstract IN33B-1174, 2008
  - SPIE Defense & Security Symposium, Optical Pattern Recognition, Apr. 2009
  - IJCAI-09 International Joint Conference on ArtificialIntelligence, Cross-media Information Access and Mining (CIAM2009), July 2009.