A hybrid object-based/pixel-based classification approach to detect geophysical phenomena

Xiang Li, Rahul Ramachandran*, Sara Graves, Sunil Movva
Information Technology and Systems Center
University of Alabama in Huntsville
rramachandran@itsc.uah.edu

Project Details

- Collaborative pilot project funded by Ames Research Center and Goddard Space Flight Center
- Other Collaborators:
  - Dave Emmitt and Steven Greco, Simpson Weather Associates
  - Robert Atlas, Goddard Space Flight Center (now Director, NOAA Atlantic Oceanographic and Meteorological Laboratory)
  - Joe Terry and Juan Carlos Jusem, SAIC
- Science goals:
  - Create a "climatology" of targets (geophysical phenomena)
- One of the Project goals:
  - *Develop an automated method to detect fronts in model data using data mining techniques*

Definition

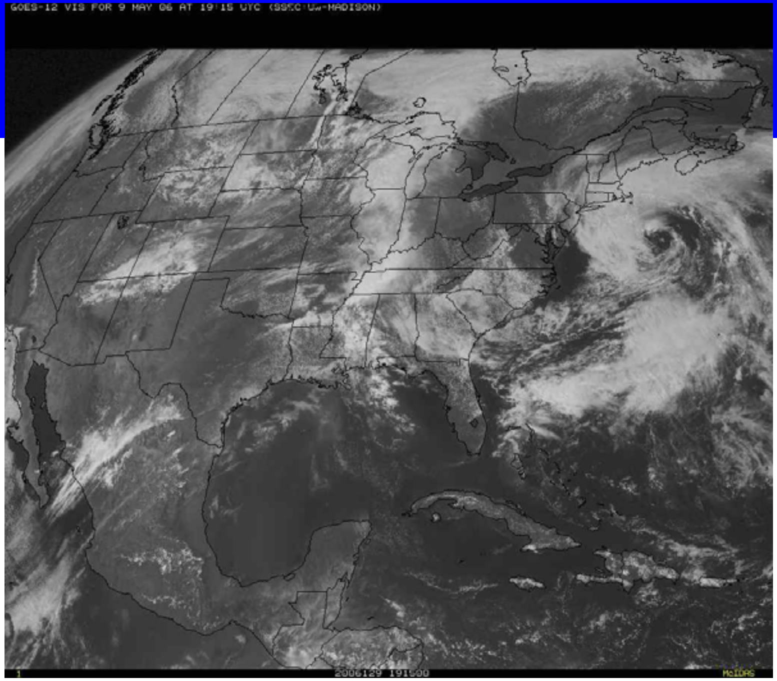
- Phenomenon
  - any state or process known through the senses rather than by intuition or reasoning, and thus is an observable event, especially something special or unusual
- Geophysical phenomenon (in the context of the geosciences data)
  - Is a significantly different region compared to the rest of the image (by spectral intensity, intensity variance, etc.)
  - Has spatial extent and may have strong shape constraint
  - Evolves, grows and perishes over time

Why do we detect phenomena?

- Studying them can improve our scientific understanding and knowledge about mechanisms that generate these phenomena.
- Timely detection is important for some geophysical phenomena which can impact human lives (severe storms, tornados, hurricanes, etc).
- Use to validate the accuracy of physical models and compare different models.

Examples

Fire and smoke



Cloud systems related to cyclone and front

Hurricane

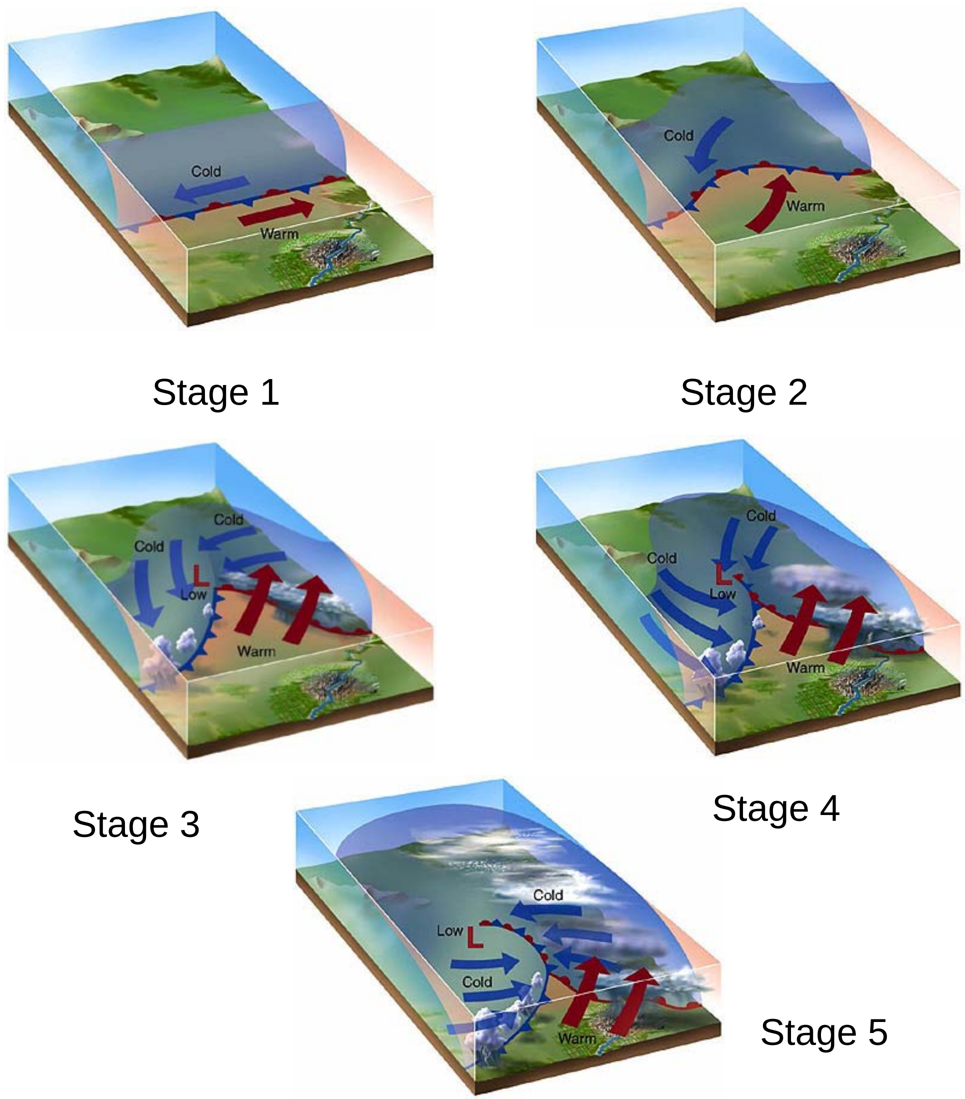
Sand storm

Approaches to detect geophysical phenomena

- Heuristic detection using domain knowledge
  - Identify rules (from domain knowledge) which can separate phenomena from background
    - Thresholding the pressure field for hurricane detection
- Use segmentation algorithms based on image statistics
  - Kapurs, Yager, Otsu, PEA
- Detection using data mining techniques
  - Use classifiers to detect

Front Detection: Background

- A front is a transition zone where two air masses meet.
- It is characterized by large temperature gradient and significant variation of wind flow patterns in transition zone.
- A front is a very complicated atmospheric phenomenon.
- It has a distinct shape but it varies depending on the type of fronts (cold/warm fronts, stationary fronts, occluded fronts) and different stages of its life time.
- Orographic effects and troughs can mimic a front
- Traditional pixel-based classification approach would not be appropriate.

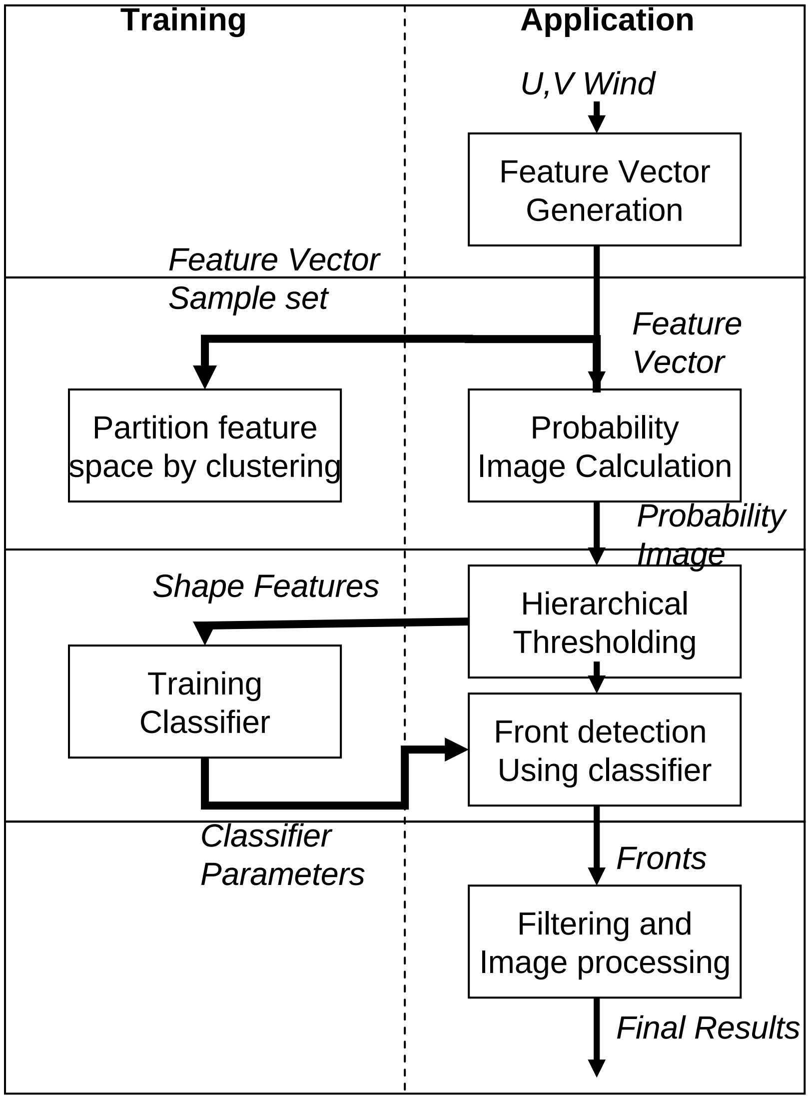


Life time of a frontal system

Diagram source: http://rst.gsfc.nasa.gov/Sect14/Sect14_1d.html

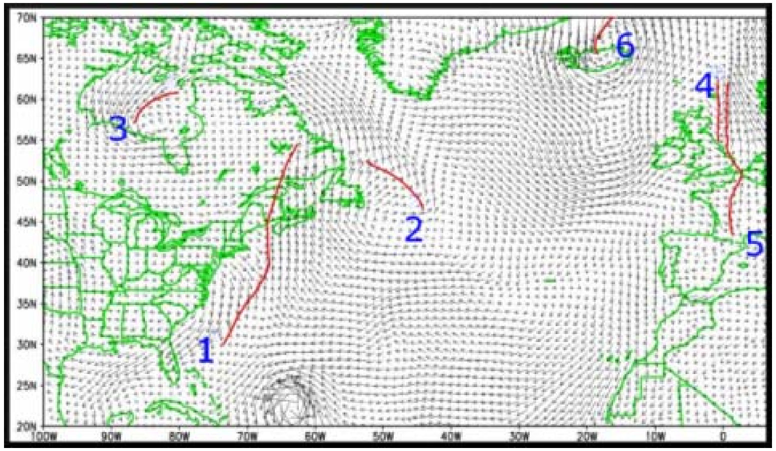
Methodology Overview

- A hybrid object-based/pixel-based classification approach used to tackle this problem.
- A soft classifier based on K-means is used to estimate the probability that a pixel belongs to a front.
- Fronts are detected at the object level by hierarchical thresholding the probability image in conjunction with a classifier
  - Classifier: Maximum Likelihood
  - Feature Vector: Region and Boundary based Shape Parameters
- The methodology is generic and can be extended to other applications which require both pixel and object level analyses.



Data and area of study

- <u>Atmospheric Model:</u> *finite volume Community Climate Model (fvCCM).*
- <u>Model resolutions:</u>
  - *horizontal:* 0.5 ° Lat x 0.625 ° Lon.
  - *vertical:* 35 levels + surface.
  - *data output:* every 6 hours
- <u>Data field used in the research:</u>
  - *surface wind fields*
  - *surface potential temperature (PT)*
- <u>Time period:</u> *9/12/1999 – 9/19/1999.*
- <u>Area of interest:</u>
  - *20°N - 70°N Latitude,*
  - *100°W - 10°E Longitude.*

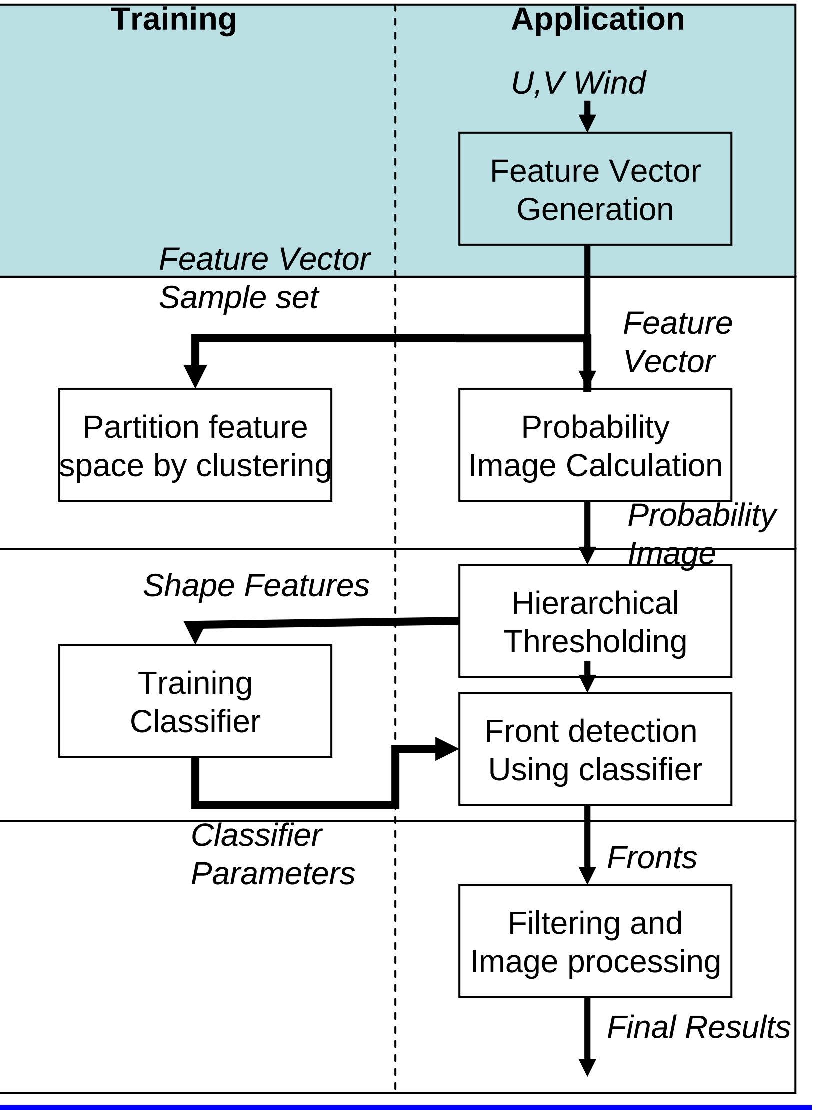


Surface wind field 09/1200Z

Locations of frontal systems were identified by domain experts as truth for time period 9/12/1999 – 9/13/1999.

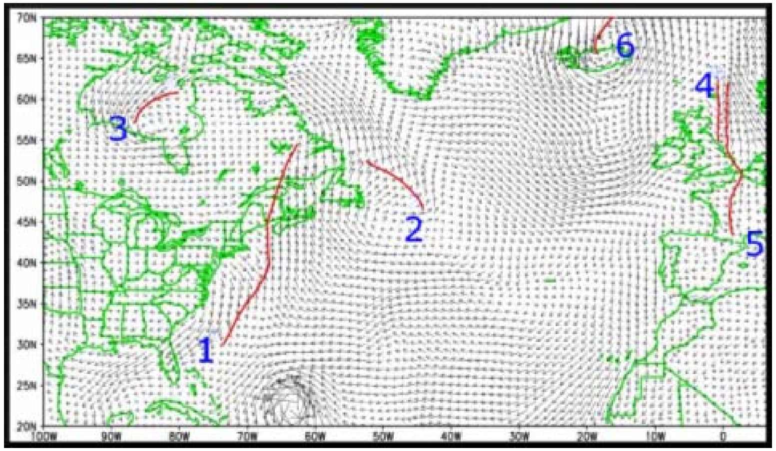
Methodology: Data Preparation

- Crucial step in any successful data mining application.
- Initially tried edge detection operators directly to wind fields for front detection but did not get good results
- Domain experts suggested using derived fields more appropriate to characterize fronts
  - Vorticity – rotation
  - Confluence – convergence
  - Wind direction variation (WDV)
  - Wind speed

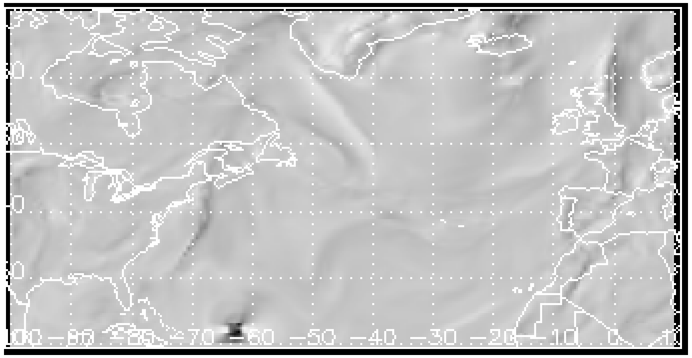


UAH The University of Alabama in Huntsville

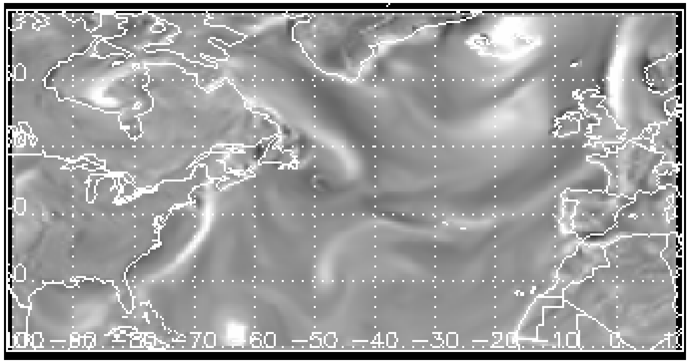
Original wind field and derived fields



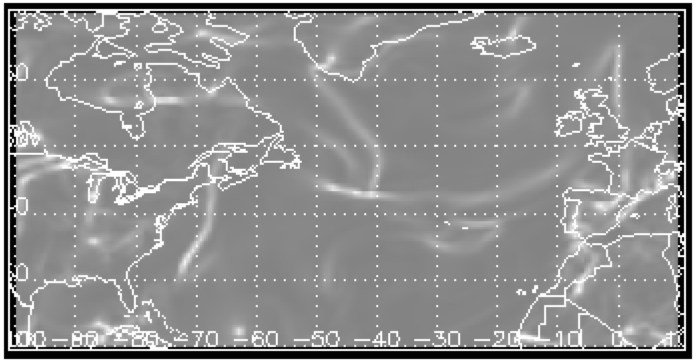
Wind field 09/12/00Z



Confluence



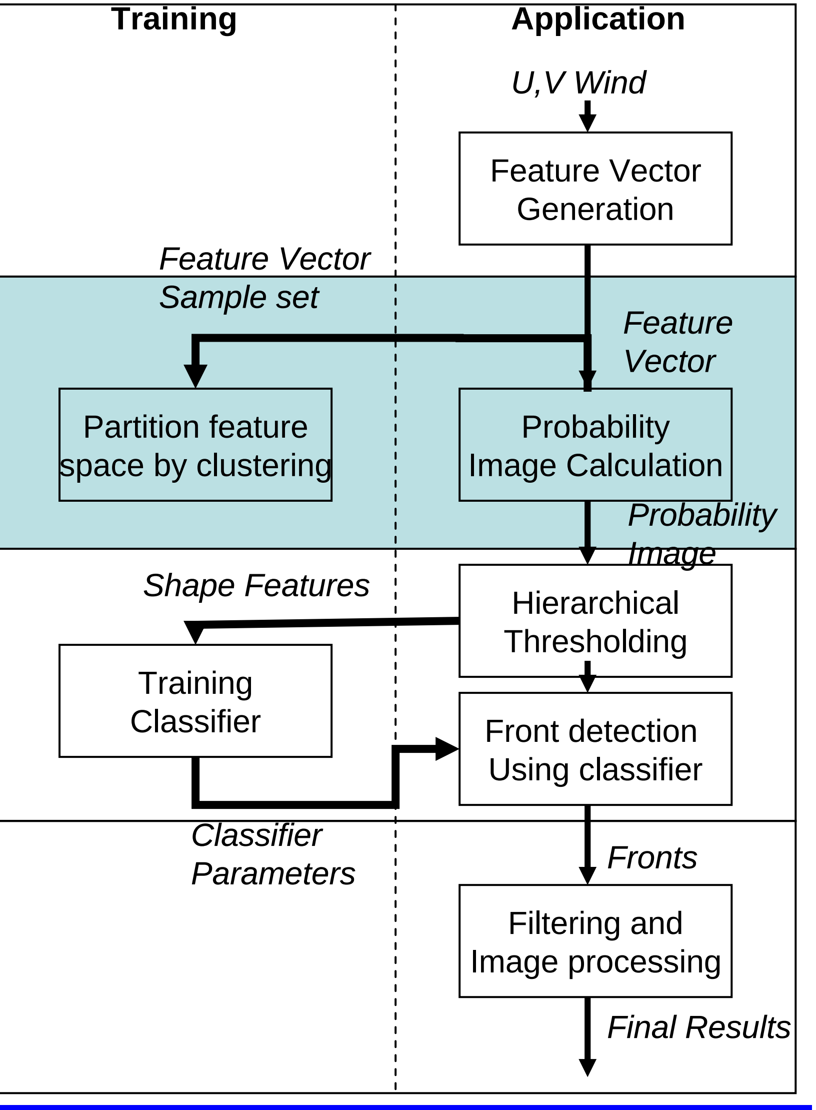
Vorticity

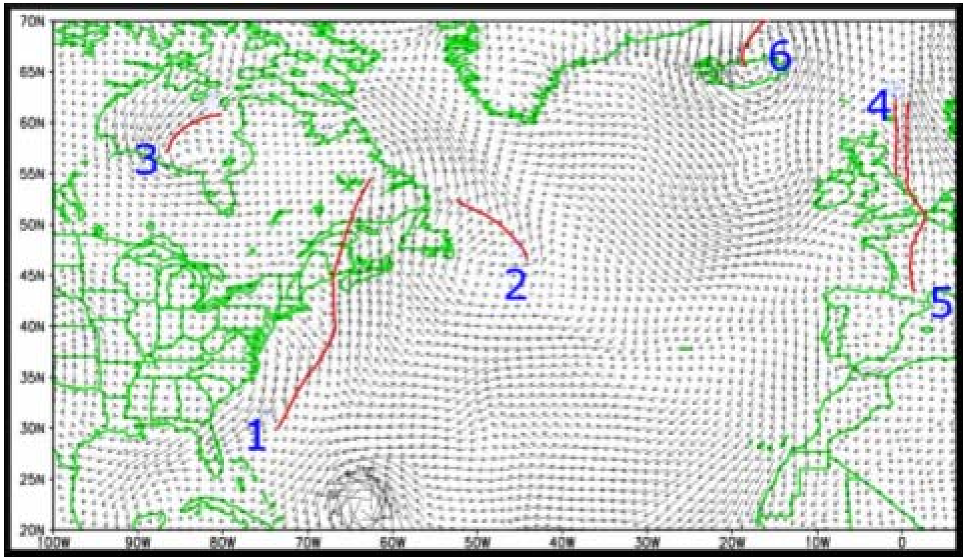


Wind direction variation (WDV)

Methodology: Pixel level processing

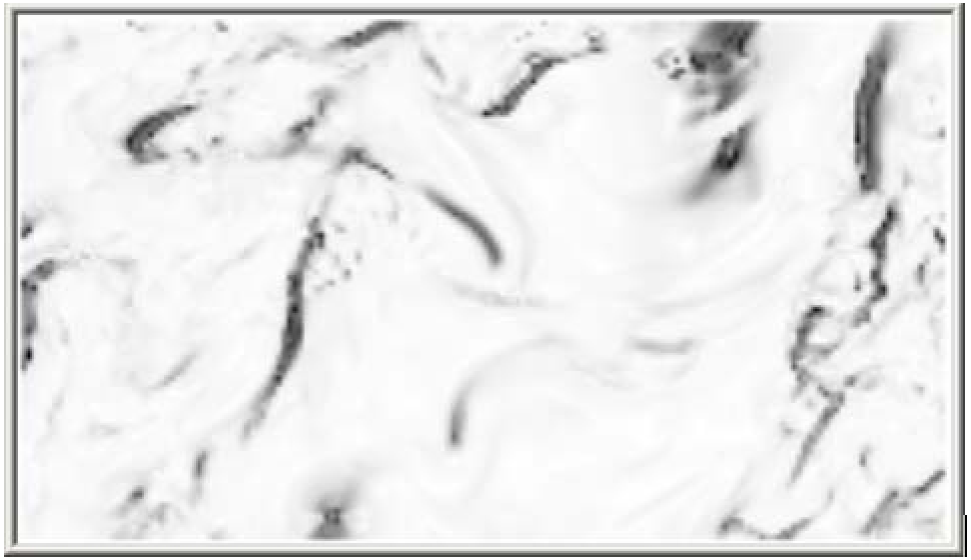
- Soft classification using K-Means clustering algorithm
  - Use the K-means to partition feature space into clusters
  - Determine clusters which represent front using domain knowledge
  - Calculate the probability based on membership function for each pixel that belongs to a front
  - Generate a front probability image





Wind field at 09/12/00Z

The darker the pixel,
the higher the probability



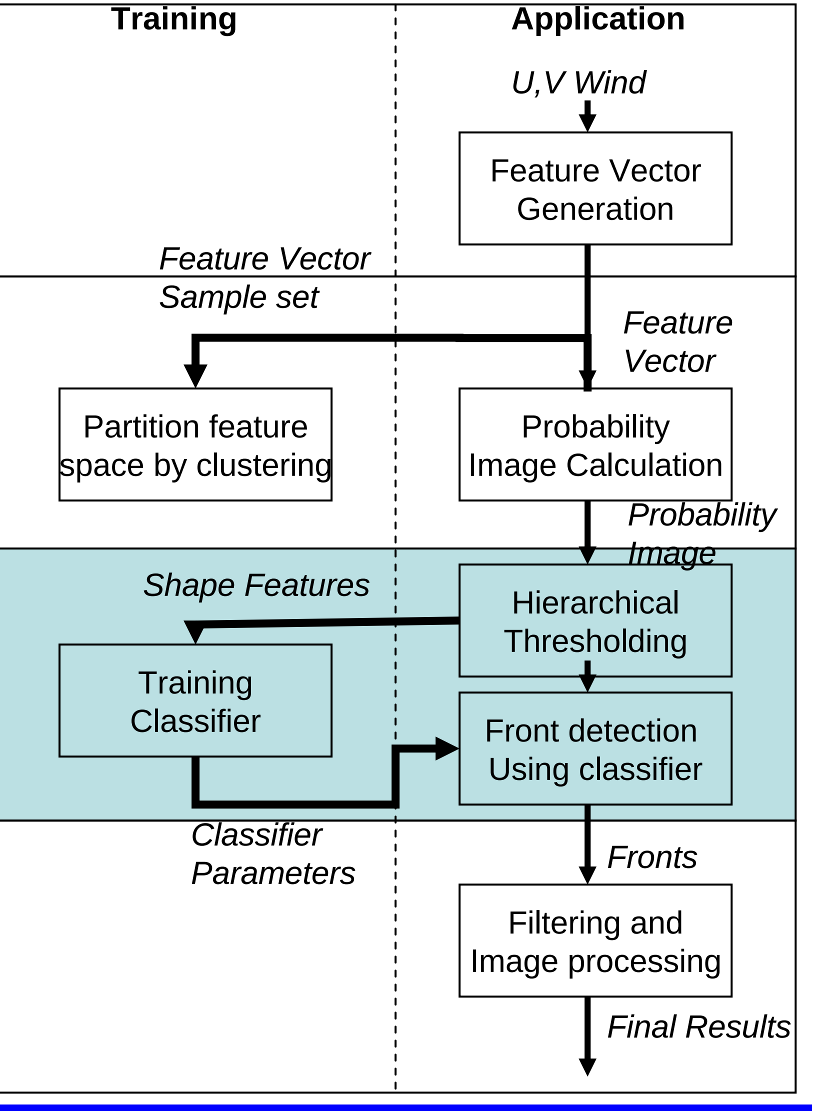
Front probability image at 09/12/00

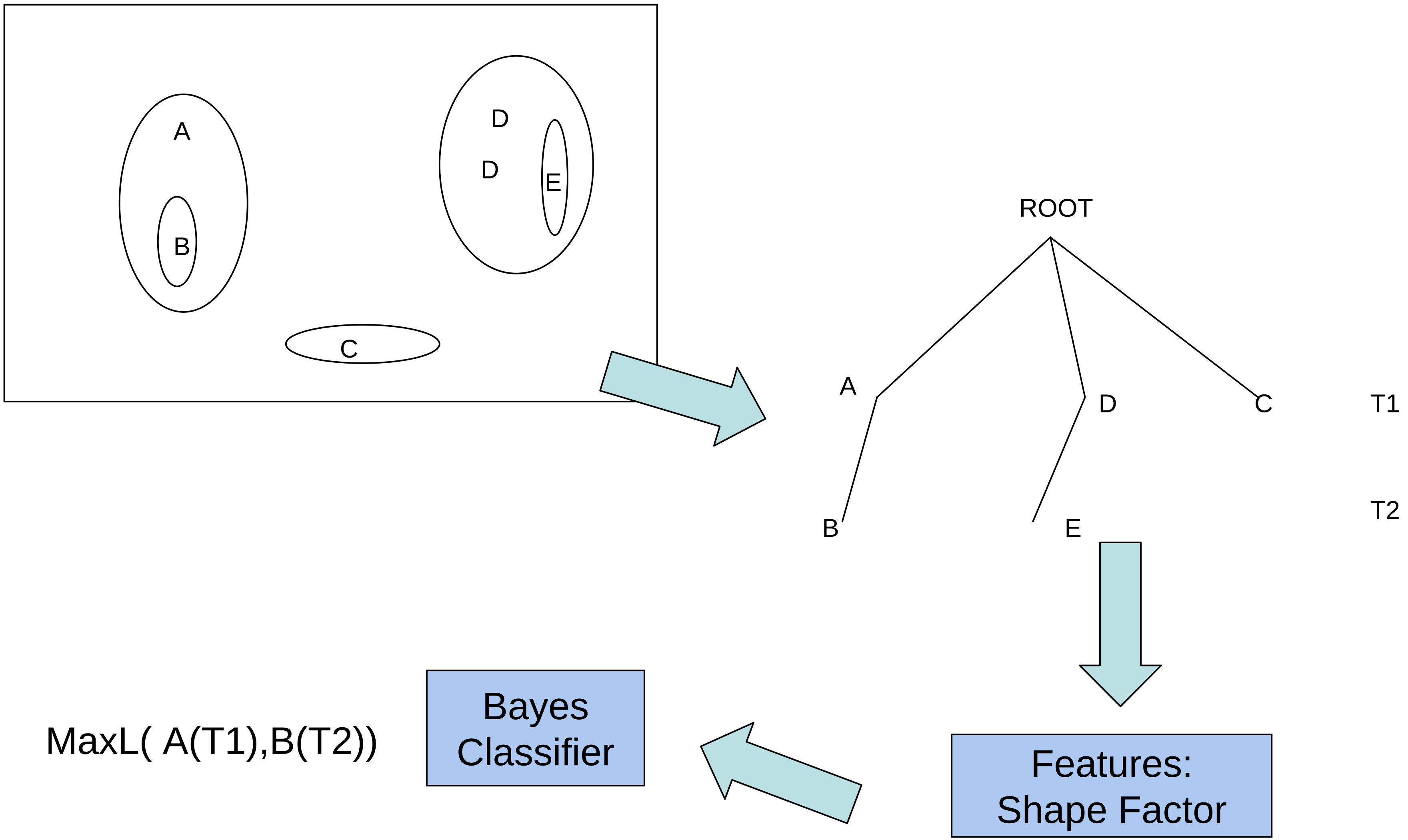
All frontal systems identified by experts corresponds to high probability values in probability image.

Probability intensities vary for each frontal system.

Methodology: Object level processing

- The probability image can be segmented to get the frontal systems using a threshold – *However* - there is no global threshold that is optimal for all fronts in an image
- Instead of global thresholding, hierarchical thresholding is used to detect fronts
- Multiple thresholds are applied to probability image.
- The optimal threshold is the one that the detected object produces maximum likelihood from a Gaussian classifier based on its shape features.

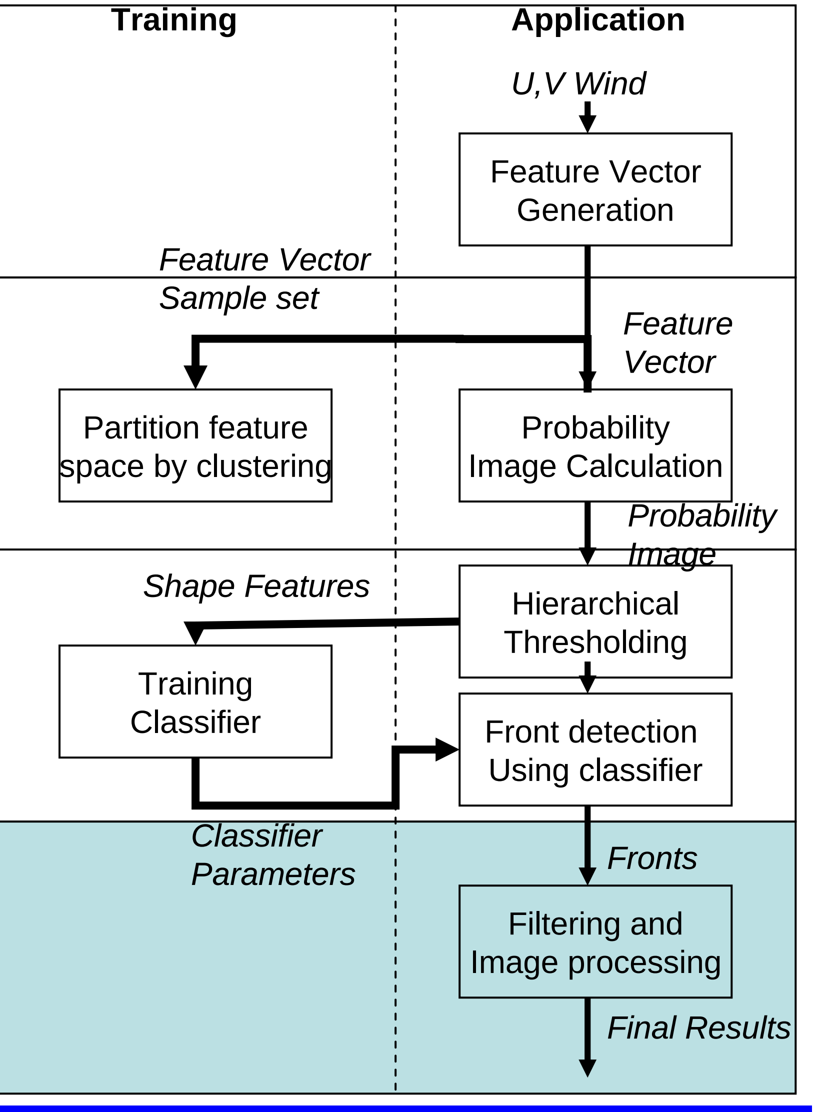




Eight shape parameters are used to characterize front shape and these parameters are derived based on region boundary, its size and its convex hull.

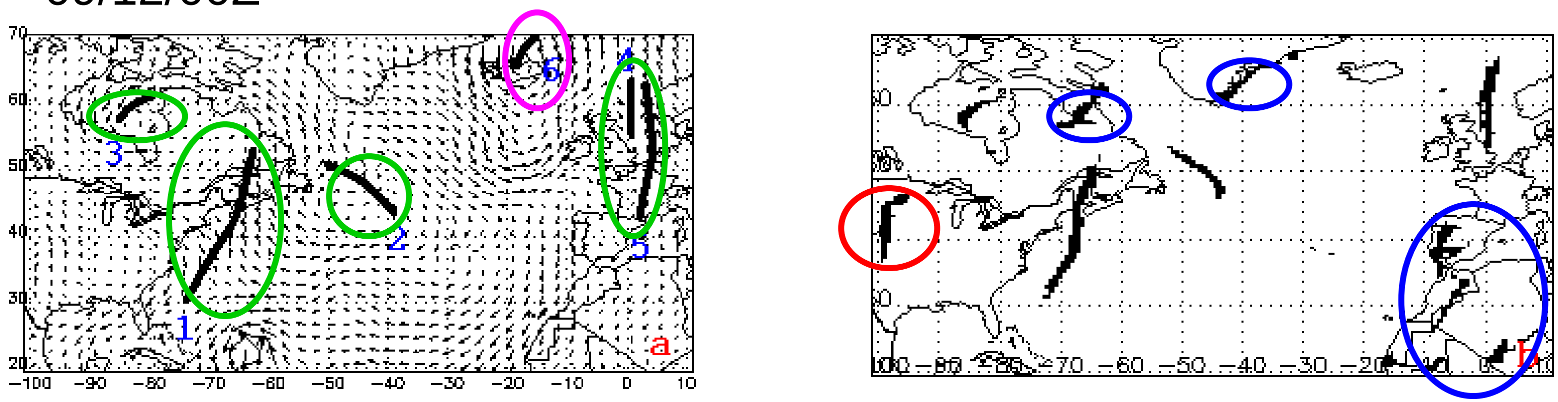
Methodology: Post processing

- Heuristics based on domain knowledge used for filtering
  - A front should have a large temperature gradient
  - The gradient will be different over water and land
- Morphological closing operation was applied to smooth and thin the detected regions

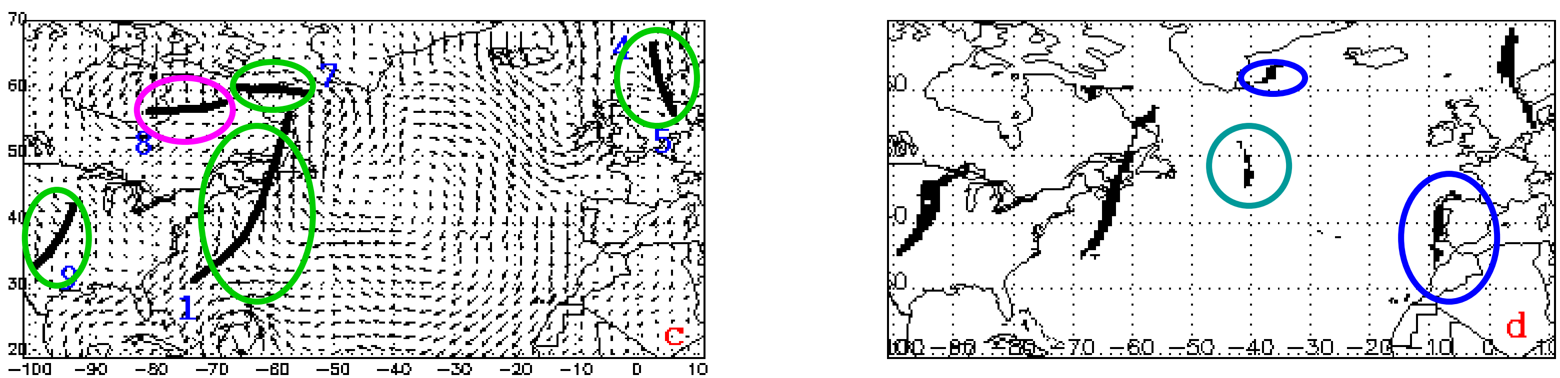


Results

09/12/00Z



09/12/12Z



Correctly detected

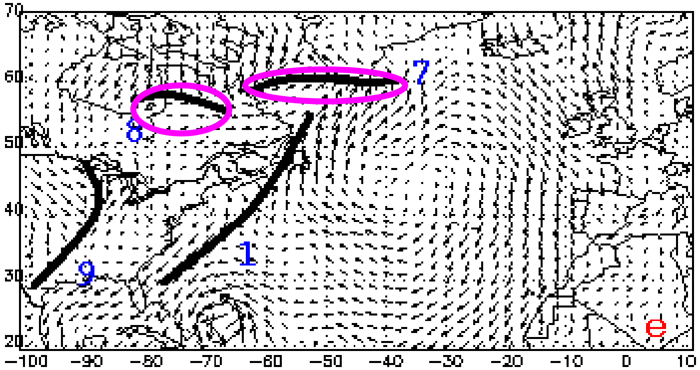
Incorrectly detected

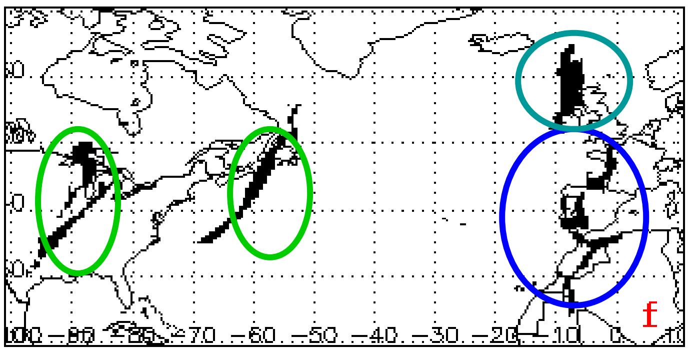
Ambiguous

Detected but not labeled by experts

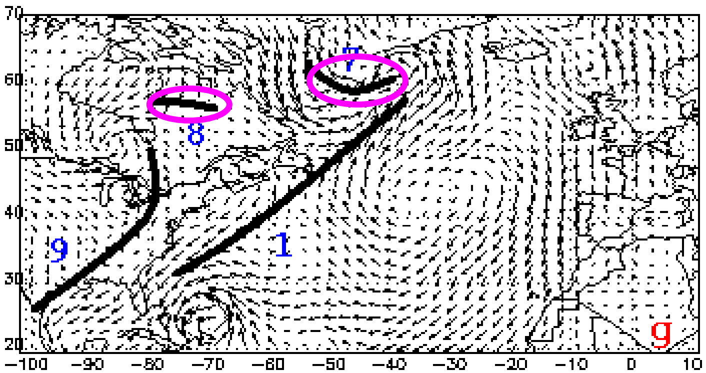
Failed to detect

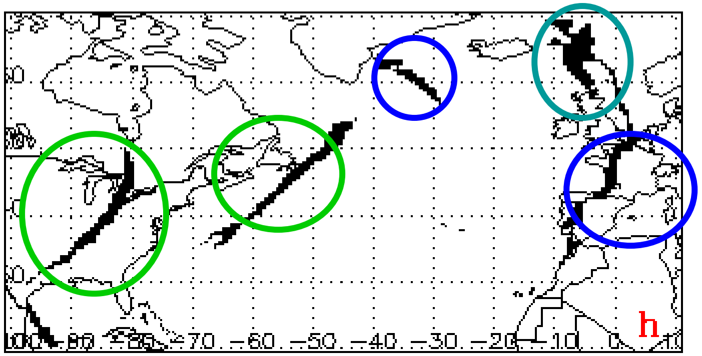
09/13/00Z





09/13/12Z





- Correctly detected (green)
- Incorrectly detected (blue)
- Ambiguous (teal)
- Detected but not labeled by experts (red)
- Fail to detect (magenta)

Analysis

- Total of 34 frontal systems were identified by the experts for the time steps 09/12/00 – 09/13/18 and were used as the truth data.
- The automated methodology detected 27 of the 34 frontal systems with the detection rate of 79.4%.
- The methodology also detected a frontal system at the early stage of its development that was not identified by the domain experts.
- The methodology also falsely identified fronts – some attributed to orographic effects

Conclusions

- An automated **hybrid object/pixel level** frontal detection method was developed using a number of data mining techniques.
- Soft classification and hierarchical thresholding are the two major components in the method
- Experimental results showed that our method identified most of the frontal systems as labeled by the domain experts.
- The methodology is generic and can be extended to other applications which require both pixel and object level analyses.
- Details: X. Li, R. Ramachandran, S. Graves and S. Movva, *A Hybrid Object-/Pixel-based Classification Approach for Shape-oriented Atmospheric Phenomena Detection*, IEEE Transaction on Geoscience and Remote Sensing (submitted 2006).