Wavelet Fuzzy Classification for Detecting and Tracking Region Outliers in Meteorological Data

Chang-Tien Lu
Dept. of Computer Science
Virginia Polytechnic Institute and State University
7054 Haycock Road, Falls Church, VA 22043
1-703-538-8373
ctlu@vt.edu

Lily R. Liang
Dept. of Computer Science
University of the District of Columbia
Washington, DC 20008
1-202-274-5086
lliang@udc.edu

ABSTRACT
In this paper, a wavelet fuzzy classification approach is proposed to detect and track region outliers in meteorological data. First wavelet transform is applied to meteorological data to bring up distinct patterns that might be hidden within the original data. Then a powerful image processing technique, edge detection with competitive fuzzy classifier, is extended to identify the boundary of region outlier. After that, to determine the center of the region outlier, the fuzzy-weighted average of the longitudes and latitudes of the boundary locations is computed. By linking the centers of the outlier regions within consecutive frames, the movement of a region outlier can be captured and traced. Experimental evaluation was conducted on a real-world meteorological data to examine the effectiveness of the proposed approach. This work will help discover interesting and implicit information for large volume of meteorological data.

Categories and Subject Descriptors
H.2.8 **[Database Applications]**: *Data mining, Spatial databases and GIS*

General Terms
Algorithms

Keywords
Outlier Detection, Spatial Data Mining, Fuzzy Classification

1. INTRODUCTION
Spatial databases have become a significant area both in academia and industry over the past decade. From satellite observation system to urban planning, geography related spatial data are widely used; there are also other spatial data, such as medical image and gene maps, which are also important and useful. Spatial data mining, as one of the main focuses of spatial database research, is the process of discovering implicit and useful spatial patterns or rules from large spatial data sets [9, 20]. Like traditional data mining, spatial data mining techniques can be classified into classification, clustering, trend analysis, and outlier detection.

Outliers are the observations differing from the remainder of the whole data set [2, 7]. Outliers are frequently treated as noise of the data sets. However, in some applications, outliers have real meaning and are essential components of the data as they reveal significant anomalous phenomena. Spatial outliers are observations that are inconsistent with the surrounding neighbors. They are frequently associated with disastrous natural events and the task of detecting spatial outlier is an essential part of spatial data mining. The challenges of spatial data mining arise from the following issues: i) Classical data mining techniques are designed to process numbers and categories; in contrast, spatial data is more complex and includes extended objects such as points, lines, and polygons; ii) Classical data mining techniques work with explicit inputs, whereas spatial predicates and attributes are often implicit; iii) Classical data mining techniques treats each input independently of other inputs, while spatial patterns often exhibit continuity and high autocorrelation among nearby features.

A data stream is an ordered sequence of data that arrive continuously and must be processed on line. Stream data differs from conventional stored relational data since data elements in the stream come continuously and change fast. The stream is unbounded in size and impossible to save in a physical media. It needs to be handled quickly to extract nearly real-time information. Data stream should be viewed as an infinite process consisting of data which continuously evolves with time [1]. The goals of any stream data mining technique are to mine patterns, to process user queries in a fast and accurate manner, and to compute statistics on data streams in real time [4].

In the research of the Atmospheric Sciences, huge amounts of spatial data have been collected continuously from both observation and modeling. Discovering useful patterns from these data streams, especially spatial outliers, has great practical value and can help weather forecast, environment monitoring, and climate analysis. In the meteorological data, spatial outliers or anomaly patterns are often associated with severe weather events. Such events usually do not happen at a single point but encompass an area. That is to say, they are usually two dimensional spatial outlier regions. Furthermore, the temporal and spatial changes of these regions are frequently associated with the variations of weather phenomena and climate patterns. To automatically extract these outlier regions is a critical issue.

GIS'04, November 12–13, 2004, Washington, DC, USA.

In this paper, we propose a wavelet fuzzy classification approach for detecting and tracking spatial outlier in meteorological data stream. For data of each time frame, wavelet transform is first performed along all latitudes. Then a powerful image processing technique, edge detection with competitive fuzzy classifier [11], is extended to detect the boundary of region outlier. After that, to determine the center of the region outlier, the fuzzy-weighted average of the longitudes and latitudes of the locations detected as boundary are computed. By linking the centers of the region outlier within each consecutive time frame, the movement of an outlier region can be effectively captured and tracked in a data stream. This work will help discover interesting and implicit information from large volume of meteorological data.

The rest of the paper is organized as follows. In Section 2, the problem is defined and related works are introduced. Our approach is described in Section 3. Section 4 shows the experimental results and Section 4 draws the conclusions and introduces future work.

2. PROBLEM DEFINITION AND RELATED WORKS

There have been many outlier detection algorithms proposed [2, 3, 8]. In previous works, outliers are usually disconnected points. But in weather and climate data, the outliers are frequently exhibited in irregular spatial forms such as regions. For the points enclosed in a region, the feature should be rather similar, while for the outside points surrounding the region, the feature would be distinctly different. Here, we define *a region outlier* as a group of adjoining points whose feature is inconsistent with that of their surrounding neighbors. In real atmosphere, the anomalies emerge at different spatial scales and may exhibit as various spatial shapes. This makes detection and tracking of these outliers from continuously arriving streams a challenging task.

The problem is to design an efficient and practical approach to detect and trace region outliers (could be in irregular shapes) in spatial data streams. Such approaches can help identify and monitor spatial anomalies such as hurricanes, forest fires, tornado, thunder storm, and other severe weather events from the continuous observation data. Figure 1 shows an example of region outlier in the water vapor distribution over the east coast of the USA, Atlantic Ocean, and the Gulf of Mexico. The grayscale of each coordinate reflects its corresponding amount of water vapor. As can be seen, located at the left portion of the image, there is a spot with much higher water vapor than its surrounding neighbors, which means a hurricane at the Gulf of Mexico. It is also clear that this spot is not a single point but a group of connected points (a region). It is a region outlier. There exists other region outliers in this figure; the number of region outliers detected will be determined by the pre-defined threshold provided by domain experts.

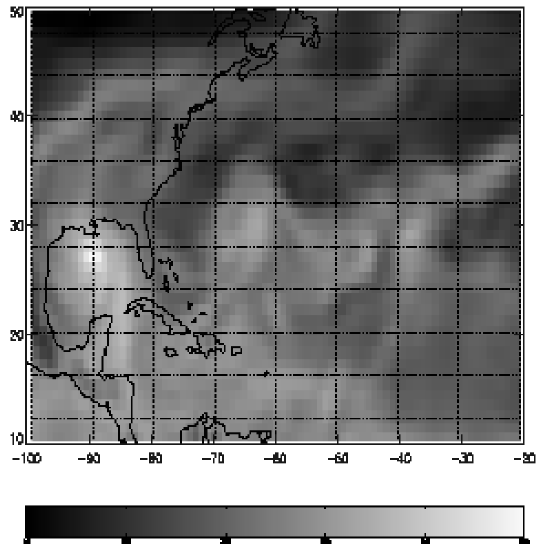

Figure 1: An example of meteorological region outlier.

Since meteorological data are usually two-dimensional spatial data and can be visualized as images, image processing techniques, e.g., edge detection, can be applied to detect meteorological region outliers. There are many different methods for edge detection [5, 6], such as Sobel filtering, Prewitt filtering, Laplacian of Gaussian filtering, moment-based operators, the Shen and Castan operator and the Canny and Deriche operator, but some common problems of these methods are their high computational cost, sensitivity to noise, anisotropy, and thick lines. Russo [16, 17], and also Russo and Ramponi [18], designed fuzzy rules for edge detection. Such rules can smooth while sharpening edges, but require a rather large rule set compared to simpler fuzzy methods [13]. Neural networks can be trained to detect edges [12] and radial basis functional link nets [15] are especially powerful for edge detection, but require training. Also, for meteorological data, feature changes are usually gradual and are more difficult to detect. That is to say, we may not be able to get the shape or coverage of the region outliers by using traditional image edge detection methods from the original data.

In recent years, wavelet analysis methods are widely used in many science and technology fields, including data mining, where it has been used jointly with clustering, classification, regression, forecasting, and data visualization [10]. Sheikholeslami, et al. developed the *WaveCluster* approach which takes advantage of the multi-scale, multi-resolution properties of wavelet analysis and clusters the wavelet-transformed spatial data in the frequency with different resolutions [19]. A wavelet-analysis-based statistic approach is introduced in [22] for detecting region outliers in meteorological data. Previous works reveal that wavelet transformation can help identify distinct patterns that might be hidden within the original data.

3. OUR APPROACH

We propose a wavelet fuzzy classification approach to detect and track region outliers in meteorological data streams. For data of each time frame, wavelet transform is first performed along all latitudes. Then a powerful image processing technique, edge detection with competitive fuzzy classifier, is extended to detect the boundary of region outlier. After that, to determine the center of the region outlier, the fuzzy-weighted average of the longitudes and latitudes of the locations detected as boundary are computed. By linking the centers of the region in each consecutive time frame, the movement of a region outlier in a data stream can be traced, and the approximate trajectory of the moving region can be captured. Details of these three steps are introduced in the following paragraphs respectively.

3.1 Wavelet transform as preparation

Wavelet transform can bring up distinct patterns that might be hidden within the original data. For meteorological data, it is preferable to decompose the original observation data into different spatial scales and treat them separately to simplify the problem and centralize the target object. Wavelet transformation provides the capability to achieve this with its multi-resolution character and the localization of variation in the frequency domain.

Compared with traditional Fourier transform, which also transfer the signal into frequency domain and separate the scales, wavelet analysis has special attractive features: I) **multi-resolution**. Wavelet analysis examines the signal at different frequencies with different resolutions; the changes of the signal at different scales may be studied with different focuses; this feature makes wavelet an effective tool to filter potential signal noises and focus on certain scales; II) **localization of the frequency**. In traditional Fourier transform, the frequency domain has no localization information, in other words, if the frequency changes with time in the signal, it is hard to distinguish which frequency happens within what time range even all the frequencies may be detected. If we need to know the exact information of a variation, the frequency and the location of a certain variation or the strength of the variation at certain location, wavelet analysis has advantages over Fourier transform.

There are continuous wavelet analysis and discrete wavelet analysis. In this paper, we use continuous wavelet analysis. For a wavelet function Ψ(t), the continuous wavelet transform of a discrete signal X_i (i=0,N-1) is defined as the convolution of X with scaled and translated Ψ, shown in equation (1):

$$W(n,s) = \sum_{i=0}^{N-1} x(i)\Psi^*\left[\frac{(i-n)\delta t}{s}\right] \quad (1)$$

where (*) indicates the complex conjugate, n is the localization of the wavelet transform and s is the scale. For the details of wavelet transform, please refer to [21].

Many functions can be used as base or mother function for wavelet analysis. We use the Mexican hat function, defined in equation (2), since it provides a better localization (spatial resolution) [22].

$$\Psi_0(\eta) = \frac{(-1)}{\sqrt{\Gamma(5/2)}}\frac{d^2}{d\eta^2}\left(e^{-\eta^2/2}\right) \quad (2)$$

We apply wavelet transform on the data along each latitude line, rather than along longitude line, since the scale in weather system is typically represented on latitude. In addition, the variations along longitude line are mostly normal patterns, such as the differences between tropics and high latitude areas, and are not the anomalous features of interest. Therefore, we focus on detecting the spatial variation along the latitude (X-axis). The wavelet transformed power indicates the strength of the variation along the latitude. The locations with high wavelet power value are the places where anomalies exist. We will concentrate on small scale weather systems such as hurricanes and tornadoes, as spatial outliers are usually small in size compared with the environment.

Wavelet power mainly represents the variation of the signal on the spatial domain. For meteorological data analysis, we should focus on the spatial variation, rather than the value of the variable. The wavelet transform provide a better description of the variation, thus making it an effective tool for pre-processing original data. Another advantage of using wavelet is its multi-scale capability. We can focus on only the scale of particular interest. For the multi-scale data such as meteorological data, this makes the complicated variation convenient to be studied.

3.2 Fuzzy classification to detect boundary

In the next step, we extend a powerful edge detection technique, competitive fuzzy edge detector (CFED) [11], to detect the boundaries of region outliers from wavelet transformed meteorological data.

An image is a two dimensional array where each element is called a pixel and represents a point in a corresponding two-dimensional space. Two-dimensional spatial data can be transformed into an image for visualization. Edge pixels in an image are defined as locations where there is a significant variation in gray level (or intensity level of color) in a fixed direction across a few pixels [7]. They are outliers in images and form curved or straight boundaries. Edge detection is by far the most common approach for detecting meaningful discontinuities in the gray level.

The wavelet transformed data are first processed with a threshold. Only data above the threshold are kept and the rest are suppressed to 0. This step cuts out low-power data items and leaves only regions of high wavelet power. To make the later parameter setting of the fuzzy classifier easier, the data are then mapped to the range of [0-255], which is the range of grayscales for an image. Then CFED is applied to detect region outlier boundary.

The CFED detects data on the region outlier boundary by fuzzy classification in the first step and applies competitive rules as a second step for the purpose of thinning the ridges around local maxima. A third step despeckles by removing single and double pixel noise specks. In the case of diffuse region outlier boundaries, it still output thin lines while most other edge detection methods result in thick lines. Compared with the widely used Canny edge detector, which also outputs thin lines, it takes much less computation and is less sensitive to noises [11].

Given an input dataset ***D***, which is the result of the thresholding and mapping process, we define directions on a 3x3 spatial neighborhood of a data item as shown in Figure 2.

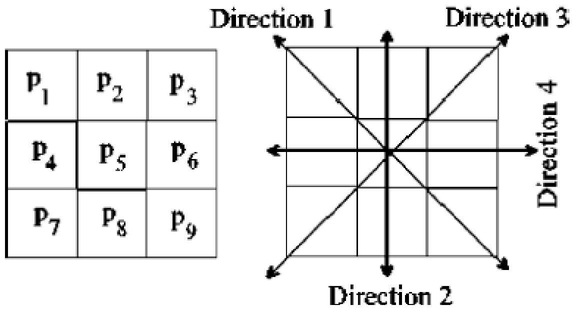

Figure 2: Spatial data and directions in a 3x3 neighborhood.

The bi-directional summed magnitude differences between data item p_5 and its neighbors are designated by d_1, d_2, d_3 and d_4 for Directions 1, 2, 3 and 4, respectively, are calculated by:

d1=|p1-p5|+|p9-p5| (Direction 1), (3a)
d2=|p2-p5|+|p8-p5| (Direction 2), (3b)
d3=|p3-p5|+|p7-p5| (Direction 3), (3c)
d4=|p4-p5|+|p6-p5| (Direction 4), (3d)

For each data item in $\boldsymbol{D}$, we compute a four-dimensional feature vector $\boldsymbol{x} = (d_1, d_2, d_3, d_4)$ of summed magnitude differences in four directions on its 3x3 spatial neighborhood. The magnitudes make each difference d_j bidirectional**.**

Data items are classified into four edge classes, a background class and a speckle edge class (a speckle is a noisy data item). Four typical neighborhood situations are used for each edge class: each directional edge neighborhood shown in Figure 3, its rotation by 180^o, and the exchange of darker and lighter pixels in each of these two cases.

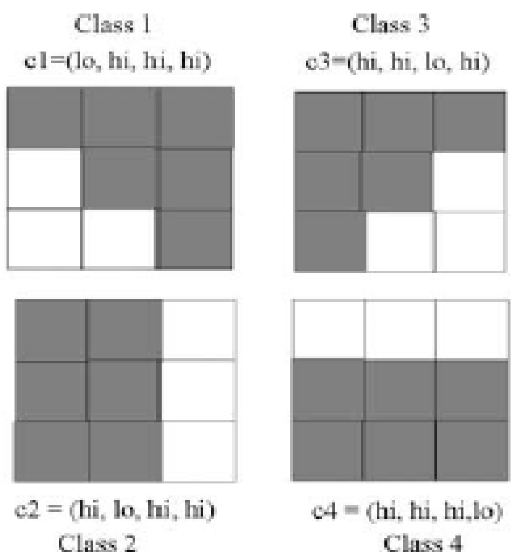

Figure 3: Edge classes.

Each set of four situations for a class has a single feature vector of summed magnitudes of differences as far as the low and high values are concerned. The background class is for any data whose neighborhood has low magnitude differences in the four directions. A speckle edge class is used for data on whose neighborhood the change magnitudes in all directions are high.

We construct six prototype vectors $c_0, \ldots, c_5$ to be the respective centers of the six classes (four edge, one background and one speckle edge classes). These centers, or prototypes, for the respective classes have component values 'lo' and 'hi' that represent low and high summed magnitude differences in the directions indicated. The parameters lo and hi are to be set by the user and depend on the data set and the sensitivity desired. These class centers for the situations, some of which are displayed in Figure 3, are listed in Table 1.

Table 1: The classes and their prototype vectors.

Class 0 (background)	c_0 = (lo, lo, lo, lo)
Class 1 (edge)	c_1 = (lo, hi, hi, hi)
Class 2 (edge)	c_2 = (hi, lo, hi, hi)
Class 3 (edge)	c_3 = (hi, hi, lo, hi)
Class 4 (edge)	c_4 = (hi, hi, hi, lo)
Class 5 (speckle edge)	c_5 = (hi, hi, hi, hi)

Figure 4 shows the fuzzy classifier architecture of CFED. The input feature vector $\boldsymbol{x} = (x_1, \ldots, x_N)$ are feeded directly to six output nodes, each of which represents a class with an extended Epanechnikov fuzzy membership [14] centered on a prototype c_j (see Eqs. (4a)–(4f)). It activates those fuzzy set membership functions, one of which will be a maximum. The output layer node with the maximum value determines the class. A data item is thus classified as one of four types of edges, a non-edge or a speckle edge.

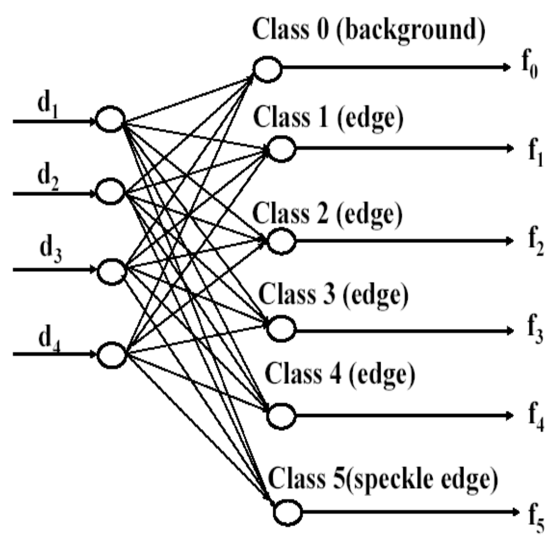

Figure 4: Edge-detection fuzzy classifier.

Class 0 (background): Max { 0, 1- $\|D\text{-}\mu_0\|2/\beta^2$} (4a)
Class 1 (Edge) : Max { 0, 1- $\|D\text{-}\mu_1\|2/\beta^2$} (4b)
Class 2 (Edge) : Max { 0, 1- $\|D\text{-}\mu_2\|2/\beta^2$} (4c)

Class 3 (Edge) : Max { 0, 1- $||D\text{-}\mu_3||2/\beta^2$} (4d)
Class 4 (Edge) : Max { 0, 1- $||D\text{-}\mu_4||2/\beta^2$} (4e)
Class 5 (Speckle edge): Max { 0, 1- $||D\text{-}\mu_5||2/\beta^2$} (4f)

To thin the edges, we next apply a competitive rule to each edge data item according to its assigned class. Each data item which is classified as an edge competes with the two data items on either side of it across the edge width. For these three data items, only the one with the largest difference magnitude is marked as boundary item. The rules for this competition are given below:

IF ***x*** is Class 0 (background)
 THEN do not mark as boundary item
IF x is Class 1 (edge)
 THEN compete d_3 with neighbor pixels in Direc. 3
  IF it wins THEN mark as boundary item
  ELSE do not mark
IF x is Class 2 (edge)
 THEN compete d_4 with neighbor pixels in Direc. 4
   IF it wins THEN mark as boundary item
   ELSE do not mark
IF x is Class 3 (edge)
 THEN compete d_1 with neighbor pixels in Direc. 1
  IF it wins THEN mark as boundary item
  ELSEdo not mark.
IF x is Class 4 (edge)
 THEN compete d_2 with neighbor pixels in Direc. 2
  IF it wins THEN mark as boundary item
  ELSE do not mark.
IF x is Class 5 (speckle edge) THEN mark as boundary item

In summary, the extended CFED model classifies data items into four different edge classes, a background class and a speckle edge class, based on their wavelet power. Competition rules are then applied accordingly to thin the detected edges. So even the region outlier has gradual wavelet power changes from the rest of the data set, the region outlier boundary can still be accurately located, which might not be achieved by most other edge detection approaches.

3.3 Fuzzy-weighted average to track center

To locate the center of the region outlier, we compute the fuzzy-weighted average of longitudes and latitudes of the locations that are marked as boundary. Fuzzy-weighted average is used make the center more representative. The movement of the region outlier is tracked by linking the centers in consecutive time frames.

A fuzzy weight is a weight determined by a fuzzy membership function, for example, the reciprocal of distances, shown in equation (5).

$$w_q = 1/D_q \qquad (5)$$

where D_q is the distance between the longitude (or latitude) of the center and the longitude (or latitude) of the q-th location marked as boundary. If the distance is relatively large, then w_q will be much smaller than if the distance is small. Average achieved with these fuzzy weights is immune to outlier vectors and more representative for densely located vectors. Since D_q can be zero we use

$$w_q = 1/[D_q + 1] \qquad (6)$$

The weights are standardized so that they sum to unity over the vector set.

$$w'_q = \frac{w_q}{\sum_{q=1}^{Q} w_q} \qquad (7)$$

Then the fuzzy-weighted average (FWA) of the longitudes (or latitudes) $\mathbf{v}_q$ of the boundary locations is computed as in equation (8).

$$\mathbf{FWA} = \mathbf{w'}_q \mathbf{v}_q \qquad (8)$$

We start with an initial center for the longitudes, which is the mean of longitudes of boundary locations. Then, i) compute weights of longitudes of boundary locations according to their distances to the current center; ii) Update the center with the fuzzy-weighted average of these longitudes. Step i and Step ii are repeated until the process converges (the longitude center makes little shift between two iterations). The final version of the longitude of the center is the longitude of the center of the region outlier. The process to obtain the latitude of the center is similar.

The center obtained by the above approach, compared with traditional methods, such as mean, median, regular average, is immune to noise in the detected boundary and is more representative for the region outlier.

4. EXPERIMENT RESULTS

In the experiment, we used NOAA/NCEP (National Centers of Environmental Prediction) global reanalysis data sets, which is a multiple-parameter data with a resolution of 1 degree by 1 degree. The data covers the whole earth and is updated 4 times a day, namely, 00AM, 06AM, 12PM, 18PM. For this particular study, we used the data of water vapor, as water vapor is a good indicator for depicting the weather system. We used data on September 18, 2003, during which hurricane Isabel landed the east coast of USA.

Hurricane Isabel formed in the central Atlantic Ocean on September 6th, 2003. It moved in a general west-northwestward direction and strengthened to a category five hurricane by 11 September. Weakening began on 16 September as the hurricane turned northwestward. Isabel made landfall on the Outer Banks of North Carolina on 18 September as a category two hurricane. Portions of eastern North Carolina and Southeastern Virginia experienced hurricane-force winds. Experiment results on hurricane Isabel demonstrate the effectiveness of our algorithms in detecting and tracking abnormal meteorological patterns.

4.1 Wavelet transform

Figure 5 shows the original global water vapor distribution between 180°W and 75°E at 0AM on September 18, 2003. In general, the tropical region is covered by high value of water vapor. We first performed Mexican hat wavelet analysis on data over all latitudes, which reveal more significant anomalous features than over longitudes, and primarily focused on the anomalies with sub-weather scales, that is, the variation of 1000km or 10 degrees at mid-latitude region. Figure 6 shows the wavelet transform power with scale index 3. As can be seen, there are some areas where the power is especially high than their surrounding neighbors. The spatial variations in these areas are prominent and they are potential region outliers. Besides, the bright hot spot center at 32°N and 72°W stands out and can be effectively identified, which corresponds to Hurricane Isabel. Notice that a high feature value does not necessarily generate a high wavelet power.

4.2 Fuzzy classification to detect boundary

In this experiment, we validated the effectiveness of the proposed fuzzy classification method on the transformed wavelet power image. Figure 7 shows the wavelet image at 0AM on September 18, 2003, with the detected region outlier boundary marked. As can be seen, the boundary of Hurricane Isabel is accurately extracted by the proposed fuzzy classification approach. Figure 8 shows another experiment result on September 18, 2003, at 18:00PM. The boundary of Hurricane Isabel is also clearly identified.

For the purpose of comparison, we also applied our approach on the original data without the wavelet transform procedure on data of 18:00PM, September 18, 2003 with the same cutoff percentage for thresholding. Then data can be mapped into the range of [0, 255], and then the fuzzy classification can be used to identify the potential outlier regions. The experiment result is shown in Figure 9. As can be observed, the boundary of Hurricane Isabel was not identified as in Figure 8. In addition, extra boundaries were detected and formed noisy output.

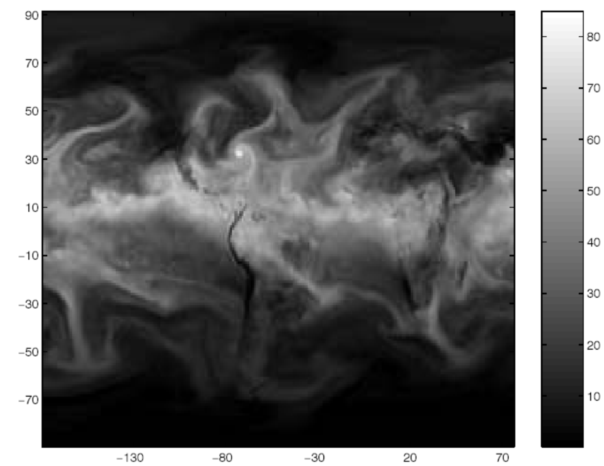

Figure 5: Global water vapor distribution at 0 AM Sep. 18, 2003.

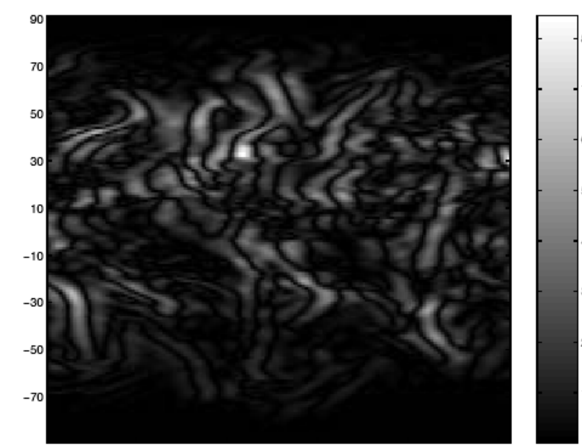

Figure 6: Wavelet power distribution at scale index 3 at 0 AM, Sep. 18, 2003.

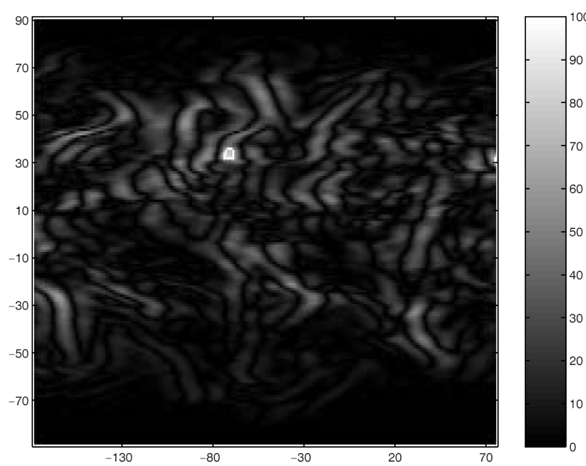

Figure 7: Wavelet power distribution at 0 AM, Sep. 18, 2003, with Hurricane Isabel identified.

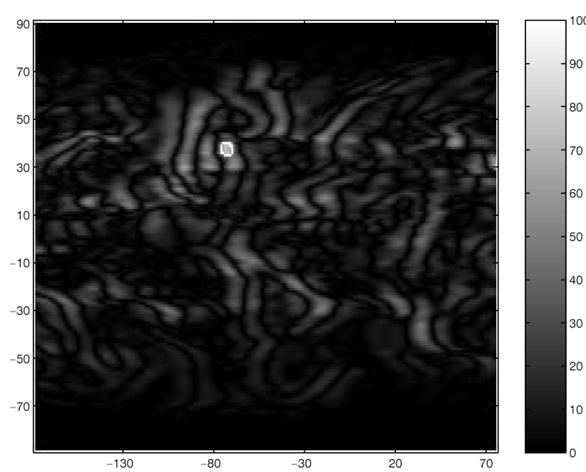

Figure 8: Wavelet power distribution at 18 PM, Sep. 18, 2003, with Hurricane Isabel identified.

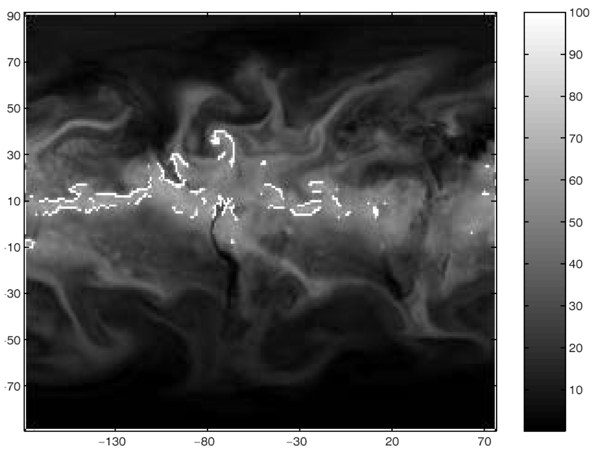

Figure 9: Detected boundary on original data at 18 PM, Sep. 18, 2003.

4.3 Tracking the center of outlier region

After the boundaries of region outliers are identified, the center of each region outlier can be computed using the proposed fuzzy-weighted average approach. We tested our method on the four consecutive wavelet data on September 18, 2003. Table 2 lists the computed centers of Hurricane Isabel from 0A.M. to 18A.M. Figure 10 shows the 3D trajectory of the center movement of four consecutive region outliers for consecutive time frames. Four regions are illustrated in this figure. The boundary of each region outlier is depicted by dotted line and their center points are connected for continuous frames, so that its moving trajectory can be observed. The trend of Hurricane Isabel can be observed as that it moves northwestward to the inner land.

Table 2: Computed center for each detected region outlier using fuzzy-weighted average

	1 (0AM)	2 (6AM)	3 (12PM)	4 (18PM)
Latitude	**32.57**	**33.29**	**34.51**	**36.42**
Longitude	**-72.45**	**-72.73**	**-73.25**	**-74.19**

5. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a wavelet fuzzy classification approach to detect and track region outliers in meteorological data. Mexican Hat wavelet transform is used to distinguish spatial region outliers. A powerful image processing technique, edge detection with competitive fuzzy classifier, is extended to detect the boundary of region outlier. The boundary of region outlier is computed to determine the representative center using the fuzzy-weighted average approach. And their trajectory can be plotted by linking these central points for consecutive frames. The experiment results demonstrate that our approaches can effectively discover anomalies corresponding to severe weather events. In the future, we are planning to study region outlier in three-dimensional spatial space with multiple attributes, such as the combination of pressure, rain fall, wind, cloud, and temperature, and to simultaneously track multiple moving region outliers for the continuous meteorological data streams.

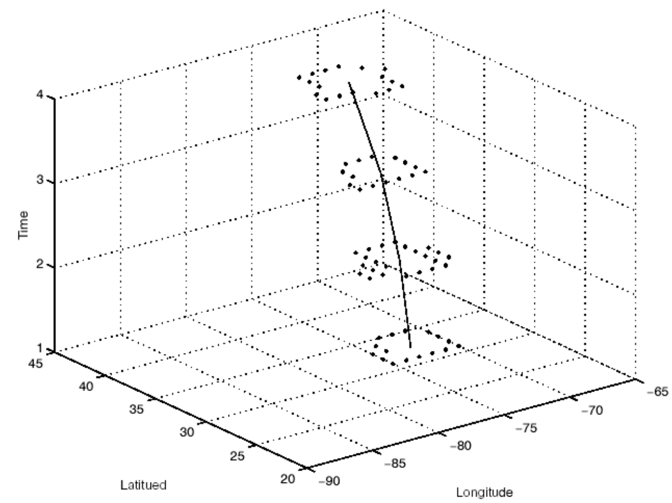

Figure 10: 3-D center trajectory for the detected region outlier within each time frame at 0 AM ~ 18 PM, September 18, 2003.

REFERENCES

[1] Aggarwal, C.C., Han, J., Wang, J., and Yu, P.S. *A Framework for Clustering Evolving Data Streams*. in *Proceedings of the 29th International Conference on Very Large Data Bases*. 2003.

[2] Barnett, V. and Lewis, T., *Outliers in Statistical Data*. 1994, New York: John Wiley.

[3] Breunig, M., Kriegel, H., Ng, R.T., and Sander, J. *OPTICS-OF: Identifying Local Outliers*. in *Proceedings of the PKDD '99, Prague, Czech Republic, Lecture Notes in Computer Science (LNAI 1704)*. 1999. Springer Verlag.

[4] Bulut, A. and Singh, A.K. *Swat: Hierarchical stream summarization in large networks*. in *Proceedings of the ICDE*. 2003.

[5] Efford, N., *Digital Image Processing*. 2000, MA: Addison-Wesley, Reading,. 164-173.

[6] Gose, E., Johnsonbaug, R., and Jost, S., *Pattern Recognition and Image Analysis*. 1996, Englewood Cliffs, NJ: Prentice-Hall. 298.

[7] Hawkins, D., *Identification of outliers*. 1980, Massachusetts: Chapman and Hall, Reading.

[8] Johnson, T., Kwok, I., and Ng., R. *Fast Computation of 2-Dimensional Depth Contours*. in *Proceedings of the the Fourth International Conference on Knowledge Discovery and Data Mining*. 1998: AAAI Press.

[9] Koperski, K., Adhikary, J., and Han, J. *Spatial data mining: Progress and challenges*. in *Proceedings of the In Workshop on Research Issues on Data Mining and Knowledge Discovery(DMKD'96)*. 1996. Montreal, Canada.

[10] Li, T., Li, Q., Zhu, S., and Ogihara, M., *A survey on wavelet applications in data mining.* SIGKDD Explorations, 2002. **4**(2): p. 49-67.

[11] Liang, L.R. and Looney, C.G. *Competitive Fuzzy Edge Detection*. in *Proceedings of the International Journal of Applied Soft Computing*. September, 2003.

[12] Looney, C.G., *Pattern Recognition Using Neural Networks*. 1997, New York: Oxford University Press.

[13] Looney, C.G., *Nonlinear rule-based convolution for refocusing.* Real Time Imaging, 2000. **6**: p. 29-37.

[14] Looney, C.G., *A Fuzzy Classifier Network with Ellipsoidal Epanechnikovs, Technical Report*. 2001, Computer Science Department, University of Nevada, Reno, NV.

[15] Looney, C.G., *Radial basis functional link nets and fuzzy reasoning.* Neurocomputing, 2002. **48**(1-4): p. 489-509.

[16] Russo, F. *A user-friendly research tool for image processing with fuzzy rules*. in *Proceedings of the Proceedings of the First IEEE International Conference on Fuzzy Systems*. 1992. San Diego.

[17] Russo, F. *A new class of fuzzy operators for image processing*. in *Proceedings of the IEEE International Conference on Neural Networks*. 1993.

[18] Russo, F. and Ramponi, G., *Fuzzy operator for sharpening of noisy images.* IEEE Electron. Lett., 1992(28): p. 1715-1717.

[19] Sheikholeslami, G., Chatterjee, S., and Zhang., A. *Wavecluster: A multi-resolution clustering approach for very large spatial databases.* in *Proceedings of the 24th VLDB Conference*. 1998.

[20] Shekhar, S. and Huang, Y. *Co-location Rules Mining: A Summary of Results*. in *Proceedings of the Spatio-temporal Symposium on Databases*. 2001.

[21] Torrence, C. and Compo, G., *A practical guide to wavelet analysis.* Bulletin of the American Meteorological Society. **79**(1): p. 61-78.

[22] Zhao, J., Lu, C.T., and Kou, Y. *Detecting Region Outliers in Meteorological Data*. in *Proceedings of the Proceedings of the 11th International Symposium on Advances in Geographic Information Systems*. 2003. New Orleans, Louisiana.