

# SKELETON-BASED TORNADO HOOK ECHO DETECTION

Hongkai Wang, Robert E. Mercer, John L. Barron

Paul Joe

Dept. of Computer Science  
The University of Western Ontario  
London, Ontario, Canada N6A 5B7  
{hwang252, mercer, barron}@csd.uwo.ca

King City Radar Station  
Meteorological Service of Canada  
Toronto, Ontario, Canada M3H 5T4  
Paul.Joe@ec.gc.ca

## ABSTRACT

We propose and evaluate a method to identify tornadoes automatically in Doppler radar imagery by detecting hook echoes, which are important signatures of tornadoes, in Doppler radar precipitation density data. Our method uses a skeleton to represent 2D storm shapes. To characterize hook echoes, we propose four shape features of skeletons: curvature, curve orientation, thickness variation, boundary proximity, and two shape properties of tornadoes: southwest localization and the ratio of storm size to model hook echo size. To evaluate the hook echo detection algorithm, the hook echoes detected in several radar datasets by the algorithm are compared to those proposed by an expert. The effectiveness of the algorithm is quantified using a Critical Success Index (CSI) analysis.

**Index Terms**— Doppler radar, precipitation density, tornado signatures, skeletons, hook echoes

## 1. INTRODUCTION

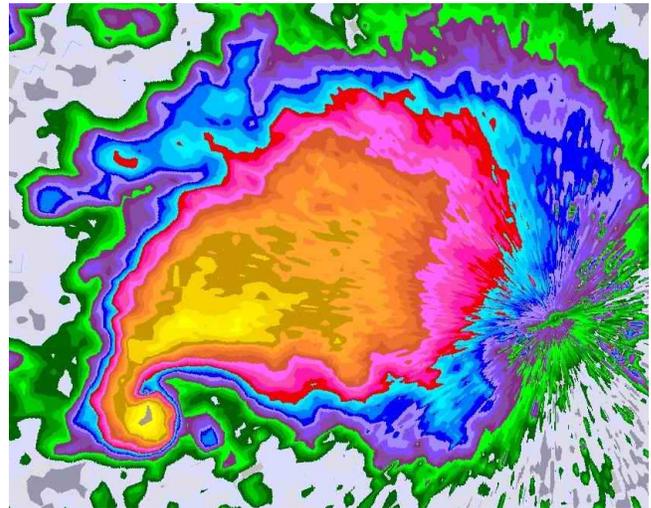
Every year tornadoes cause considerable damage and loss of life in North America and around the world. Automatic tornado detection is of great interest to meteorologists for the nowcasting/forecasting of tornadoes from Doppler radar images. The **hook echo** is an important signature of tornado existence, having been of interest to meteorologists for more than 50 years (Markowski's recent review [1] indicates the first observation of a hook echo in conventional radar was on April 9<sup>th</sup>, 1953 by the Illinois State Water Survey). A classic hook echo can be seen as the yellow and orange hook in the bottom left corner of the image in Figure 1, a precipitation density image taken from the dataset of radar images of the Oklahoma tornado outbreak on May 3<sup>rd</sup> 1999.

## 2. RELATED WORK

3D Doppler radar provides weather information through its precipitation density, radial velocity and spectrum width (velocity variance) data. The NexRad Level II datasets that we are working with were collected by the WSR88D (Weather Surveillance Radar C 1988 Doppler) weather radar system. See [2] for details.

### 2.1. LITERATURE SURVEY

In this paper, three tasks are addressed: (1) how to represent the 2D shape of a storm; (2) how to model salient features of



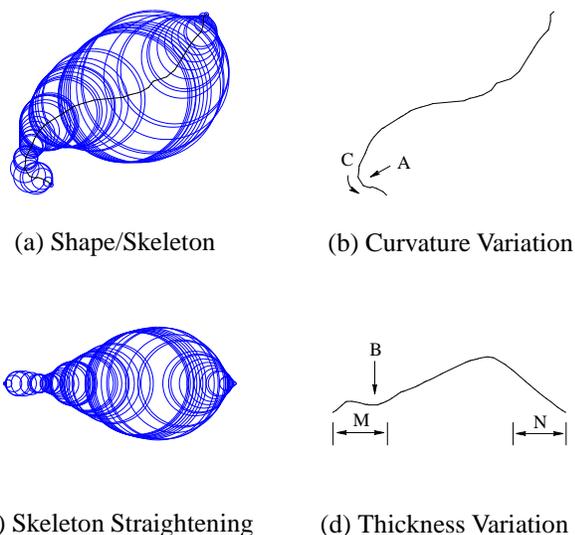
**Fig. 1.** A classic hook echo (colored yellow and orange), with the color mapping of the precipitation density value shown on the right hand side. The tornado associated with this echo was part of the May 3<sup>rd</sup> 1999 Oklahoma severe weather storm.

a hook echo constrained by this representation; and (3) how to use this model to find hook echoes that correspond to tornadoes in radar data. We use skeletons to perform these tasks. 2D skeletonization is a process of transforming a 2D object into a 1D curve representation (See Figure 2a).

We use the Hierarchic Voronoi Skeleton (HVS) algorithm [3] in our work. The HVS algorithm extracts skeletons of shapes in multi scales, based on the salience of different skeleton branches and forms a skeleton pyramid in a coarse to fine structure. An example of the highest level skeleton computed for a storm is shown in Figure 2a. The HVS algorithm is ideal for our application because it is tunable for extracting the most salient skeleton backbone of a storm shape.

### 2.2. HOOK ECHO DETECTION RESEARCH

Current NEXRAD radar systems have some automatic algorithms to identify circulation signatures in Doppler radar observations. The WSR-88D MDA/TDA algorithms use the change in radial velocity to identify circulations and shears



**Fig. 2.** (a) The skeleton of an actual storm shape and its disk reconstruction (the disk centered at each medial axis point is the largest disk contained in the shape), (b) the curvature variation along the skeleton (A indicates a high curvature location and C indicates the hook’s orientation), (c) straightening the skeleton to eliminate its curvature property and only show its thickness variation property and (d) the thickness variation along the skeleton (B indicates the fat-thin-fat location).

[4]. The linear least squared derivative technique (LLSD) computes rational shear and divergence from the radial velocity observations using a least square fit of the data [5]. Multi-Doppler based 2D wind field analysis computes the 2D wind vectors from the radial velocity observations from two or more radars with sufficient viewing angle differences [6]. The Dual Polarimetric debris signature technique uses the information from polarimetric observations that debris within tornadic events have a lower cross-correlation coefficient and a differential reflectivity near zero [7]. A serious flaw in using radial velocity profiles to detect tornadoes is that the velocity measurements may be aliased because the maximum velocity magnitude that can be measured is about 45km/hr while the wind speeds in a tornado are typically 160km/hr or more.

### 3. HOOK ECHO MODELLING METHODS

In this section, we describe our skeleton-based hook echo detection approach. The basic steps of our detection algorithm are described below.

#### 3.1. STORM SEGMENTATION

The first step is to segment the storms in the Doppler radar reflectivity images. Since the HVS algorithm uses binary input, we need to: (1) bilinearly interpolate the raw precipitation density data at elevation 0 into smooth connected image data, (2) threshold the image into storms using a precipitation density value of 35dBZ (as suggested by Johnson et al. [8]) (3) eliminate unwanted artifacts (i.e. small shapes) using

two passes of  $3 \times 3$  median filtering and (4) fill “holes” in the shapes using a floodfilling based algorithm.

#### 3.2. SKELETONIZATION

We use the C source code<sup>1</sup> for HVS skeletonization [3]. We tuned this algorithm to capture the backbone and the hook-like shapes near the storm boundary. We use only the highest level (coarsest) skeleton available from a hierarchy of computed skeletons. After the skeletonization, storm shapes are converted to polylines to capture the topological information of the storms. A skeleton is composed of one or more chains of skeleton nodes (a skeleton may exhibit a bifurcation). Each skeleton node is denoted by its  $x$  and  $y$  coordinates and radius value (maximal radii of all disks centered at this point that are constrained to be completely inside the shape boundary).

#### 3.3. BASIC FEATURES AND PROPERTIES

The basic skeletal features and tornado properties that we detect are:

1. **Curvature:** the curvature of the hook skeleton has a large absolute value. As shown in Figures 2a and 2b, along the skeleton branch, the position of the hook (indicated as A) has a relatively high curvature value.
2. **Orientation:** the orientation of the hooks can help to improve the detection accuracy. In Figure 2b, the rotating arrow labelled C indicates the orientation of this hook as counterclockwise (from the inner side to the endpoint). Most tornadoes rotate counterclockwise in the northern hemisphere.
3. **Thickness Variation:** a hook skeleton has a distinctive fat-thin-fat (bottle-neck) thickness variation along the hook. This distinctive radius variation occurs along the skeleton in close proximity to the region of high skeleton curvature. Figure 2c shows the result of straightening the curved skeleton in Figure 2a. This straightening suppresses the curvature property and only the thickness variation along the skeleton is illustrated. In Figure 2d, we can see the position of a hook (indicated as B) which exhibits the fat-thin-fat property.
4. **Boundary Proximity:** the hook is near (touching or almost touching) the boundaries of the protruding parts of the storm shapes, and these protruding parts can be well captured by skeletons. As shown in Figure 2d, if a certain distance threshold value for being near such boundaries is set, for example, in the range M and N marked in the figure, then a hook can only be detected in those ranges.
5. **Southwest Localization:** Most hook echoes display their hook shapes in the southwest direction of the storm in the northern hemisphere.

<sup>1</sup><http://www.cs.sunysb.edu/~algorithm/algorithm/skeleton/distrib/>

6. **Storm Size:** a hook only happens in relatively big storms. We use the length of skeleton (the maximum arclength distance marked on the skeleton distance map) to estimate the actual storm size.

### 3.4. DISTANCE METRIC

We introduce a distance metric on skeletons: ‘Arc Length Distance’. This distance metric measures the distance from one node to another node using the accumulated sum of length of line segments along the path.

### 3.5. COMPUTATION OF CURVATURE, ORIENTATION AND SOUTHWEST LOCALIZATION

The computation of curvature and orientation are done in the same process. We evaluate the curvature value for each skeleton node using a local window based method. The algorithm we use to compute curvature is a  $2^{nd}$  order polynomial interpolation based method due to [9]. For a skeleton node of interest, Node  $b$ , we first find the two skeleton nodes, Nodes  $a$ ,  $c$  on each side along the skeleton piece that have the same arclength distance to Node  $b$  (this usually involves an interpolation calculation as explained below.) And then we use these three nodes to evaluate the curvature for the location of Node  $b$ .

Since there is usually no node that has the exact distance compatible with the parameter **window size** to the node of interest, we interpolate intermediate nodes along the skeleton which have the exact distance to the centered node. Then we fit a  $2^{nd}$  order polynomial to these three nodes (the centered node plus the two interpolated end nodes on each side) to measure the curvature [9]. We use a threshold on this value to judge whether a candidate location has enough curvature or not to be accepted or rejected as a high curvature point.

When computing the orientation, we use the same three nodes, Nodes  $a$ ,  $b$ , and  $c$  used when estimating the curvature value. First, we construct two vectors: one is from one end node to the centered node while the other is from the centered node to the other end node. Then we compute the cross vector product of these two vectors. The sign of the result indicates the orientation. We always compute orientation using skeleton nodes going from the storm boundary to its interior [2].

The southwest criterion is also computed using the two end nodes of the local window, as described above. We construct a vector, which is from one end node to the other end node along this direction. For southwest protruding hooks, this angle should be in the range  $[\pi, 3/2\pi]$ , that is, lying in the third quadrant.

### 3.6. THICKNESS VARIATION

The location of a fat-thin-fat region (see Figure 2d) corresponds to a local minimum. We compute the derivatives of this radius function,  $f(x)$ , to locate fat-thin-fat positions in the following manner. Suppose that the function  $f(x)$  can be differentiated twice and that  $x = a$  is a critical point ( $f'(a) =$

0). Then, if  $f''(a) < 0$  the critical point  $a$  is a local maximum, and if  $f''(a) > 0$  the critical point  $a$  is a local minimum. If  $f''(a) = 0$  then we can draw no conclusions. We fit a quadratic function of the form  $f(x) = ax^2 + bx + c$  to  $N$  nodes and solve (in the least squares sense) for  $a$ ,  $b$  and  $c$ . Then  $f'(x) = 2ax + b$  and  $f''(x) = 2a$ . These values allow the fat-thin-fat criterion to be evaluated at each  $x$ .

### 3.7. BOUNDARY PROXIMITY AND STORM SIZE

For boundary proximity, we only need to threshold the distance value for each node using 2 thresholds so that the node is near an endpoint but not too close. ‘Storm size’ is used to rule out small-sized storms that are very unlikely to contain hooks. We use the skeleton distance map to estimate the storm size for this feature by finding the node in the distance map that is furthest from the endpoints. Then this distance value can be used as an estimate of the skeleton size (i.e. the storm size).

## 4. EXPERIMENTAL RESULTS

We use three datasets: KTLX19990503 from May 3<sup>rd</sup>, 1999 Oklahoma City, KTLX20030508 from May 8<sup>th</sup>, 2003 Oklahoma City and KUEX20030622 from Grand Island/Hastings. For each dataset, only the lowest elevation was used. Our meteorologist, Paul Joe, marked hooks as either ‘well-formed’ (definite) or ‘marginal’ (potentially) to provide groundtruth.

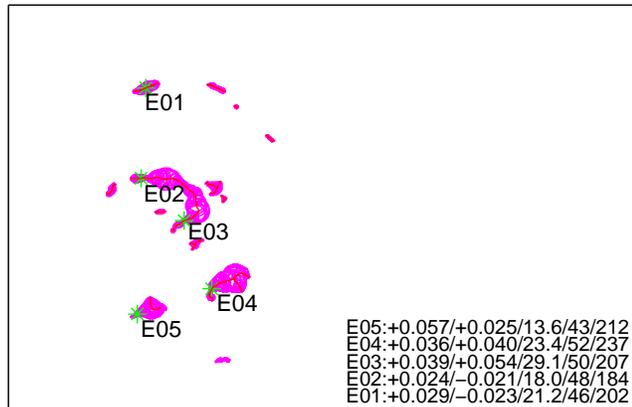
The six criteria introduced above are used to detect hook echo locations under an AND condition. Thresholds were chosen heuristically based on empirical observations [2]. We compare detection results of our algorithm with groundtruth and classify the results into three categories: **hits**: our detection matches with the groundtruth; **misses**: our groundtruth indicates it is a hook while our detection algorithm indicates it is not; and **false alarms**: groundtruth indicates it is not a hook, while our detection result indicates it is.

The results for dataset KTLX19990503 (235621e0) are shown in Figure 3. Hook echo detection results are printed on the skeletons and relevant numbers are printed in the legends of the figures. For each detected location, its curvature, orientation, thickness variation, boundary proximity, reflectivity and direction angle values are printed in the legend (lower right corner of the image). The format of a legend item is:  $E_x: y/z/m/n/p$  where  $x$  is the detected hook echo number,  $y$  is the absolute value of the curvature value, the sign of  $y$  shows the orientation (+ for counterclockwise and – for clockwise),  $z$  is the thickness variation value,  $m$  is the boundary proximity value,  $n$  is the reflectivity value and  $p$  is the direction angle value (in degrees).

We use the common statistical analysis CSI method, which is a quantitative verification based on Contingency Tables and Associated Scores. The forecast (our detection results) and observed (groundtruth from our meteorologist) are classified in a table. From the contingency table we can compute the **False Alarm Ratio** (FAR) [number of false alarms divided by number of positive forecasts – the fraction of predicted

Datasets	Well-formed Hooks Detected	Marginal Hooks Detected	False Alarms	POD Score	FAR Score	CSI Score
KTLX19990503 $1^{st}$ half, 43 frames	57/66=86.4%	30/46=65.2%	68	77.7%	43.9%	48.3%
KTLX19990503 $2^{nd}$ half, 43 frames	45/60=75.0%	42/73=57.5%	202	65.4%	69.9%	26.0%
KTLX19990503 whole, 86 frames	102/126=81.0%	72/119=60.5%	270	71.0%	60.8%	33.8%
KTLX20030508 whole, 43 Frames	21/23=91.3%	31/40=77.5%	41	82.5%	44.1%	50.0%
KUEX20030622 whole, 182 frames	40/63=63.5%	88/198=44.4%	255	49.0%	66.6%	24.8%
Total (for all datasets)	163/212=76.9%	191/357=53.5%	566	62.2%	61.5%	31.2%

**Table 1. Resulting scores for each radar dataset for our detection algorithm.**



**Fig. 3. Experimental results for radar reflectivity images March 3<sup>rd</sup>, 1999 data, 23 hours, 56 minutes and 21 seconds for elevation 0.**

hooks that do not occur], the **Probability of Detection (POD)** [number of hits divided by number of observed events – the fraction of observed events that were correctly forecast] and the **Critical Success Index (CSI)** [number of hits divided by number of hits, misses and false alarms – how well the forecasted events correspond to the observed events].

We give the experimental results for the three datasets we have studied in Table 1. These results show that our algorithm can detect well-formed hook echoes properly but for marginal hooks, the detection ratio is lower than that of well-formed hooks. This is reasonable, since marginal hooks don't have as clear a hook shape as well-formed hooks. Note the POD, FAR and CSI scores are calculated for the union of well-formed and marginal hooks to show the overall performance of our algorithm.

We split the KTLX19990503 dataset in half and analyze each set of detection results separately. As shown by the first and second row entries in Table 1, such differences exist even in the same dataset. By further analyzing the radar images, we found that our algorithm prefers clean and clear radar images which contain well separated individual storm blobs, rather than 'messy' and loosely connected storm blobs.

## 5. CONCLUSIONS

We presented an automatic tornado detection algorithm based on detecting tornado hook echoes via skeletons. We devised a number of basic features of hook echoes: curvature, ori-

entation, boundary proximity, thickness variation, southwest localization and storm size. These features work well for the hook echo modelling task. Future work includes incorporating radial velocity and spectrum width data in our algorithm, and using supervised learning to tune the algorithm parameters automatically.

## 6. REFERENCES

- [1] P. M. Markowski, "Hook echoes and rear-flank downdrafts: A review," *Monthly Weather Review*, vol. 130, no. 4, pp. 852–876, 2001.
- [2] H. Wang, "Skeleton-based hook echo detection in doppler radar precipitation density imagery," M.S. thesis, The University of Western Ontario, December 2005.
- [3] R. L. Ogniewicz and O. Kuebler, "Hierarchic voronoi skeletons," *Pattern Recognition*, vol. 28, no. 3, pp. 343–359, 1995.
- [4] E. D. Mitchell, S. V. Vasiloff, G. J. Stumpf, M. D. Eilts, A. Witt, J. T. Johnson, and K. W. Thomas, "The national severe storms laboratory tornado detection algorithm," *Weather and Forecasting*, vol. 13, pp. 352–366, 1998.
- [5] E. D. Mitchell and K. E. Elmore, "A technique for identifying regions of high shear associated with mesocyclones and tornadic vortex signatures," in *14<sup>th</sup> Intl. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, 1998, pp. 312–315.
- [6] V. Lakshmanan, T. Smith, K. Hondl, G. J. Stumpf, and A. Witt, "A real-time, three dimensional, rapidly updating, heterogeneous radar merger technique for reflectivity, velocity and derived products," *Weather and Forecasting*, vol. 21, pp. 802–823, 2005.
- [7] K. A. Scharfenberg, D. J. Miller, T. J. Schuur, P. T. Schlafetter, S. E. Giangrande, V. M. Melnikov, D. W. Burgess, Jr. D. L. Andra, M. P. Foster, and J. M. Krause, "The joint polarization experiment: Polarimetric radar in forecasting and warning decision-making," *Weather and Forecasting*, vol. 20, pp. 775–788, 2005.
- [8] J. T. Johnson, P. L. Mackeen, A. Witt, E. D. Mitchell, G. J. Stumpf, M. D. Eilts, and K. W. Thomas, "The storm cell identification and tracking algorithm: an enhanced wsr-88d algorithm," *Weather and Forecasting*, vol. 13, no. 2, pp. 263–276, 1998.
- [9] M. Marji, *On the Detection of Dominant Points on Digital Planar Curves*, Ph.D. thesis, Wayne State University, Detroit, Michigan, 2003.