Localization Methods of Weighted Centroid of dBZ on Weather-Radar Echo Maps in Vector Format

Xue-tao Yu1, 2, Xiao-ping Rui3, Feng Li1, 2
1State Key Laboratory of Coal Resources and Safe Mining, China University of Mining & Technology, Beijing, China.
2College of Geoscience and Surveying Engineering, China University of Mining & Technology, Beijing.
3College of Resources and Environment, Graduate University of Chinese Academy of Sciences, Beijing.
*Corresponding author, e-mail: ruixp@yahoo.com.cn

Abstract
Fast generation of weather-radar echo maps in vector format and accurate localization of weighted centroid of dBZ (dBZ stands for decibels of reflectivity of a radar signal reflected off a remote object) are the basis of studying the characteristic tracking algorithms which are based on the vector echoes. The authors principally studied the approach to generating the vector echo map, and discussed the localization methods of weighted centroid of dBZ on vector echo maps. First, based on the traditional calculation method on raster echo data, some new localization methods of weighted centroid of dBZ on vector echo data were proposed by considering the weights of features' area and distance from their location to radar center. Second, taking the base reflectivity products of CINRAD/SA weather radar in Meizhou city of China as data sources, they illustrated the storage structure of this type of echo data and studied the drawing mode of changing this type of data into vector format files under the polar coordinate system in detail. Third, using the same vector echo maps created by the above method, the weighted centroid of the same area was calculated by the above localization methods. In the end, Compared with the calculated value of the same area by traditional method which is based on raster echo maps, the three new calculated results and the sources of error were analyzed in detail and two conclusions were drawn: the echo's precision in vector format is much higher than that in raster format and it is more accurate to take the features' area and distance to radar center as weights during the calculation of weighted centroid of dBZ on echo maps in vector format.

Keywords: base reflectivity product, radar echo data in vector format, weighted centroid of dBZ

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction
The research of identifying, tracking and forecasting storms automatically has been carried out for half a century by using weather radar, and many theories and approaches have been developed. These approaches can be divided into three classes: Persistence, Cross Correlation and Centroid Tracking. And the last approach is the most common. It obtains the motion vectors of the same storm cell from two consecutive radar echo images and thus gets the mobility information and evolution characteristics of the storm cell [1-5]. Therefore, as one of the important characteristics of storm cell, weighted centroid of dBZ's accurate localization is the first problem to identify, track and forecast storms. At the present time most of the radar echo data used in nowcasting is the raster echo images. As the generation of raster echo images needs the coordinate conversion, the precision of this type of data will be lower to some extent. While the vector echo maps with higher precision are easy to implement the accurate calculation of the length and area, block and track the storms without coordinate conversion. Therefore, it has great guiding significance for the research of radar warning and forecast to study the methods of generating vector echo maps and how to locate the weighted centroid of dBZ based on these maps accurately.

Concerning the study about locating the weighted centroid of the base reflectivity products of CINRAD as data sources, Chen et al [6] researched the format of the base reflectivity products based on the data documentation for NexRad level III which is released by national climatic data center of America, while they didn’t make a further research on the geo-visualization of this type of data. Zhang et al [7] made researches on the extrapolation of the
squall line echo band by the method of tracking echo centroid based on the PPI data in raster format. Qiao et al [8] studied the principle of centroid method of radar echo extrapolation and made an application by using the raw data of CINRAD. Li et al [9] described the calculation method of the center of gravity of radar image. But none of them made a further research on the generation of vector echo and the localization method of weighted centroid of dBZ on these maps. Based on the traditional localization method on raster echo data, some new localization methods of weighted centroid of dBZ on vector echo maps are proposed in this paper. In the end, the authors tested these localization methods based on the vector echo maps and analyzed the results and error sources.

2. Fundamental Principle

As one of the most important characteristics of the radar echoes, weighted centroid of dBZ’s accurate calculation could improve the accuracy of cell centroid algorithms. Before the calculation of weighted centroid on raster echoes, the authors assume the starting point of the radar image locates at the lower left corner, and mark its coordinates by (mapX, mapY). They also assume the size of a pixel is (cellSizeX, cellSizeY) and pixel counts in X and Y directions are cx and cy, and also define the pixel value in row i column j on the raster image as Pij. Then the traditional calculation formula of the weighted centroid of dBZ in raster radar image is [7, 9, 10].

\[
W . x = m a p X + c e l l S i z e X \times \frac{\sum_{j=1}^{cy} \sum_{i=1}^{cx} P_{ij} \times i}{\sum_{j=1}^{cy} \sum_{i=1}^{cx} P_{ij}} \\
W . y = m a p Y + c e l l S i z e Y \times \frac{\sum_{j=1}^{cy} \sum_{i=1}^{cx} P_{ij} \times j}{\sum_{j=1}^{cy} \sum_{i=1}^{cx} P_{ij}}
\] (1)

Except the raster radar echoes, the vector radar echoes studied by Rui et al [11] and Yu et al [12] exist in actual applications too. As the generation of vector echoes from the raw radar data doesn’t need the coordinate conversion, this type of data has higher precision. It displays the echoes by recording the coordinates of triangular or quadrangles, so its locations are significant while the attributes are hidden. Besides, vector echoes have their own characteristics of vector format, such as high graphic display quality, high precision, and they are convenient to carry out the network analysis and search analysis. Therefore, vector radar echoes have some specified scope of applications, such as the identification, tracking and extrapolation based on them. During the study of centroid tracking algorithms for identifying and tracking the vector radar echoes, accurate localization of weighted centroid is the first problem for this type of tracking algorithms.

As each feature’s size and dBZ value is not equal on echo maps, the features weight can’t be considered equal to each other. The calculation formula of the weighted centroid of dBZ above is no longer suitable for this type of data. Furthermore, it is well known that the features which are more close to the radar center have higher accuracy. Therefore, the authors propose to calculate the weighted centroid of dBZ by selecting the features’ area and distance to the radar center as the weights.

\[
W . x = \frac{\sum_{i=1}^{n} F_i . x \times F_i . a r e a \times F_i . v a l u e}{\sum_{i=1}^{n} F_i . a r e a \times F_i . v a l u e} \\
W . y = \frac{\sum_{i=1}^{n} F_i . y \times F_i . a r e a \times F_i . v a l u e}{\sum_{i=1}^{n} F_i . a r e a \times F_i . v a l u e}
\] (2)
The authors assume there are \( n \) features in the custom tracking area, mark feature \( i \) by \( F_i \); its area by \( F_i.area \); its dBZ value by \( F_i.value \); its coordinate of centroid by \((F_i.x, F_i.y)\). They also define the scan radius of radar as \( R \), and the coordinate of radar center as \((Cx, Cy)\). If the weight of distance from the features to the radar center is not considered, the calculation formula of the weighted centroid of dBZ in tracking area of vector echo maps is as formula (2).

The famous weighting function of distance from one single cell to the radar center is the exponent weighting function \( w_e \) and the cressman weighting function \( w_c \) which were proposed by Xiao et al [13, 14]. Exponent weighting function \( w_e \) can be written as

\[
w_e = e^{-\frac{r^2}{R_1^2}},
\]

where \( R_1 \) is the appropriate distance scale and \( R_1 = 230 \)km which is equal to the value of the scan radius of radar in this paper and \( r \) is the distance from the grid cell to the radar center in this function. Cressman weighting function \( w_c \) is defined as

\[
 w_c = \begin{cases} 
 (R_c^2 - r^2)/(R_c^2 + r^2) & r \leq R_c \\
 0 & r > R_c 
\end{cases},
\]

where \( R_c \) is the radius of influence and \( R_c = 230 \)km which is equal to the value of the scan radius of radar in this paper and \( r \) is the distance from the grid cell to the radar center.

If the function \( w_e \) is considered too, the calculation formula of the weighted centroid of dBZ in tracking area of vector echo maps is given by

\[
 C_{e}.x = \frac{\sum_{i=1}^{n} F_i.x \times F_i.area \times F_i.value \times w_i}{\sum_{i=1}^{n} F_i.area \times F_i.value \times w_i},
\]

\[
 C_{e}.y = \frac{\sum_{i=1}^{n} F_i.y \times F_i.area \times F_i.value \times w_i}{\sum_{i=1}^{n} F_i.area \times F_i.value \times w_i}.
\]

\( w_i \) is the distance weighting function of feature \( i \) in this formula, and its expression is defined as

\[
 w_i = e^{-\frac{(cx-F_i.x)^2 + (cy-F_i.y)^2}{R_i^2}}
\]

If the function \( w_c \) is considered too, the calculation formula of the weighted centroid of dBZ in tracking area of vector echo maps is written as

\[
 C_{c}.x = \frac{\sum_{i=1}^{n} F_i.x \times F_i.area \times F_i.value \times \frac{R_2^2 - r_i^2}{R_2^2 + r_i^2}}{\sum_{i=1}^{n} F_i.area \times F_i.value \times \frac{R_2^2 - r_i^2}{R_2^2 + r_i^2}},
\]

\[
 C_{c}.y = \frac{\sum_{i=1}^{n} F_i.y \times F_i.area \times F_i.value \times \frac{R_2^2 - r_i^2}{R_2^2 + r_i^2}}{\sum_{i=1}^{n} F_i.area \times F_i.value \times \frac{R_2^2 - r_i^2}{R_2^2 + r_i^2}}.
\]

\( r_i^2 \) is the square of distance from feature \( i \) to the radar center in this formula, and its expression is the following formula (8)
For two echoes of different scan time, the moving speed of echoes is equal to the position changes divided by the time difference. As is known to all, this method is operable under the precondition that the echoes are in the view of radar. In that case, when the radar echoes just move into or out of the view, there will be errors for this calculation of weighted centroid of dBZ. Hence, the human-computer interaction should be used. The users select the same echoes from two maps of different scan time by mouse, and then calculate the moving speed and direction of the weighted centroid [9].

3. Generation of Vector Echoes

For testing and explaining the localization methods of weighted centroid on vector echo data, this paper selects the base reflectivity products of Doppler radar as data sources. However, this kind of data’s storage structure is so complex that it is very difficult to extract useful information from it. Normally, it’s only opened by PUP software to display the concerned data. Consequently, it is necessary to know the storage principle and structure of this type of data first of all.

3.1. Introduction of Base Reflectivity Products

The base reflectivity products are composed of message header block, product description block and product symbology block. This file of base reflectivity product adopts an encoding mode that upper bytes are front and lower bytes are behind, which is contrary to the present operation system’s mode.

The message header block occupies 18 bytes, including message code, date of message, time of message, source ID, destination ID, number of blocks, and so on. The product description block holds 102 bytes, storing latitude and longitude of radar, height of radar, product code, operational mode, volume coverage pattern, sequence number, volume scan number, volume scan date, volume scan start time, generation date of product, generation time of product, and so on. The product symbology block has more complicated storage structure, containing block code, length of data block, number of layers and other information of data layers. Also, data layers preserve length of layer and radial data packets, while length of data block indicates byte count of the whole block and length of layer signifies byte count of radial data packets. Figure 1 shows the block’s storage structure in detail. For radial data packets, this packets store packet code, index of first range bin, number of range bins, I and J center of sweep, scale factor, number of radials as well as radial data. In addition, each radial is composed of radial start angle, radial angle delta, run codes, color levels and so on. The sketch map of radial data packet’s storage structure is shown as Figure 2. The identification code of radial data packet is 0xAF1F where 0x denotes hexadecimal number. Radial data is consist of NumberOfRadials sectors and each sector is made by 2×NumberOfRLE little fan rings [6].

![Figure 1. Storage Structure of Product Symbology Block](image1)

![Figure 2. Storage Structure of Radial Data Packet](image2)
The authors successfully analyzed the echo data with C# after comprehending the storage structure of products. At first, a nested structure is founded depending on storage structure of each data block and all items of echo data are temporarily stored according to the size of each item. Binary Reader class is used to read the three blocks of echo data in turn. Additionally, because of the difference of data encoding pattern, the bytes’ order of short integer or integer type of data needs to be reversed. Then, these byte arrays are transferred into short integer or integer with BitConverter class according to the different types of data.

### 3.2. Generation of Vector Echo Maps

Firstly, the authors will simply introduce the scanning strategies of Doppler radar before drawing vector echoes. The volume coverage pattern includes VCP11, VCP21, VCP31 and VCP32. It scans 14 elevation angles within 5 minutes under the VCP11 pattern, 9 elevation angles within 6 minutes under the VCP21 pattern, and 5 elevation angles within 10 minutes under the VCP31 and VCP32 patterns. Detailed scanning elevation angles under the four patterns can refer to particular books [15].

Owing to the original storage coordinate of echo data is polar coordinate, polar coordinate is used to draw echo map in order to maintain figure's authenticity and reduce amount of calculation. Multiple data exhibits quite a number of color level values are equal to zero, which displays current region is in no echo state, so there is no need to draw the echoes. In this case, the number of drawn polygons is enormously lowered. For the drawing colors, they can refer to the color band of base reflectivity product in China Weather Website.

Prior to drawing radial data, it is necessary to calculate the longitude incremental \( (I_{kilometerdx}) \) and latitude incremental \( (I_{kilometerdy}) \) per one kilometer near the radar center on the latitude and longitude line. Due to tiny coordinate’s variation nearby the center point, the authors consider the longitude and latitude coordinates are uniformly changed. The calculating formulas of longitude incremental and latitude incremental are given by formula (9)

\[
\begin{align*}
I_{kilometerdx} &= 360 / \left[ \cos\left( centerY \times \pi / 180 \right) \times 40075.67 \right] \\
I_{kilometerdy} &= 360 / 40009
\end{align*}
\]

(9)

where \( centerY \) expresses latitude value at the radar center.

After that, longitude incremental \( (AddXPerFactor) \) and latitude incremental \( (AddYPerFactor) \) of one unit range bin are also needed to be calculated by formula (10)

\[
\begin{align*}
AddXPerFactor &= ScaleFactor / 1000 \times I_{kilometerdx} \\
AddYPerFactor &= ScaleFactor / 1000 \times I_{kilometerdy}
\end{align*}
\]

(10)

where \( ScaleFactor \) means the count of pixels composing one range bin. Afterwards, echoes are drawn in a way of double loops which external loop traverses all radial data while inner loop traverses all range bins in current radial data and sector or fan ring will be approximately replaced with triangle or polygon [11,12]. Taking \( j \) range bin of \( i \) scanning radial data for example (\( j \neq 1 \), i.e., current radial range bin is not the first), they calculate the coordinates of four points on two neighboring lines in this radial data which are shown as Figure 3, such as C, D, K and J point. Here, the calculating formula of D point’ coordinate is given by formula (11)

\[
\begin{align*}
D_x &= centerX + \sin(RadialPacket.radials[i].StartAngle) \times r2 \times AddXPerFactor \\
D_y &= centerY + \cos(RadialPacket.radials[i].StartAngle) \times r2 \times AddYPerFactor
\end{align*}
\]

(11)

where \( centerX \) and \( centerY \) present longitude and latitude coordinate at radar center, \( RadialPacket.radials[i] \). Start Angle signifies the starting azimuth of \( i \) scanning radial data, \( r2 \) denotes the distance from \( j \) range bin to radar center. Other points’ calculating formulas are similar to D point’s one. Quadrangle can be drawn and its color level value is assigned in the attribute table after acquiring all points’ coordinates. In this way, other features could be drawn and hence the radar vector map under polar coordinate system is accomplished.
Nevertheless, created files in shp format wouldn’t directly reflect echo’s colors in the GIS software if the colors aren’t reset. The echo files are visualized and rendered by the method of unique value in accordance with the color level value from attribute table on ArcGIS Engine platform. As shown in Figure 4, base reflectivity map of Meizhou in vector format is drawn and rendered according to this method. The precision of retrieved echoes in this way is higher than those in raster way. That is mainly because original radial data is directly converted and outputted in the same coordinate system without any error from coordinate conversion. Moreover, this method’s efficiency is also very high, for example, three seconds are merely consumed in the whole process containing generation, output and display.

4. Localization of the Weighted Centroid

As there are lots of features in the vector map created by the method above, some of them may be not completely closed. So the geometries of features should be checked and repaired before other operations. One of the solutions is to call the Geoprocessor and RepairGeometry classes in ArcGIS Engine. Firstly, the FeatureClass of vector echo layer is assigned into the in_features of RepairGeometry, and then call the Execute command of Geoprocessor whose parameter is the RepairGeometry class defined above and execute this command in the end. After these steps, all the features will be repaired and completely closed.

The vector echo layer after repaired can meet the request for weighted centroid of dBZ’s calculation in the custom tracking area or whole map. For the calculation in custom tracking area, the authors design such an operation that users loop a pre-analysis polygon on the vector echo map by mouse, clip and save it as a new layer. For clipping the vector layer, IBasicGeoprocessor and ITopologicalOperator methods are utilized in ArcGIS Engine. The first method is based on ITable type, and the operation result is type of IFeatureClass which can be directly exported as one layer. After repeated experiments, the authors find that this method has low accuracy and big error may occur in the attribute of features as a result of the operation. So
they tested the second method which is based on IGeometry type. By the way of traversing all the features, they export and save the features which are within the border of polygon entirely and the intersections of features and the polygon’s border. The clipped vector echo layer by the second method is shown in Figure 5.

For the weighted centroid of dBZ’s calculation on the vector echo layer after clipped, if the weight of distance from the features to the radar center is not considered, the calculation method of the centroid is required to traverse all the features and compute the accumulation by formula (2). If the weight of distance is considered, the method is similar by using formula (5) and (7).

5. Results and Discussions

The authors calculate the coordinates of weighted centroid of dBZ by using formula (2), formula (5) and formula (7) for the same selected area in Figure 5. The results are shown in Table 1. They transform the same product data into a raster image in img format according to the grid conversion algorithm proposed by Yu et al [16], and the raster image is shown in Figure 6. The coordinates of weighted centroid of the same area on this image are also calculated by formula (1). They are (116.66660925 °E, 24.20033073 °N). Besides, the coordinates of the location of Doppler radar in Meizhou city are (115.993 °E, 24.255 °N). After the calculation, the distance from the weighted centroid on this image to the radar center is 68651.4458 meters.

<table>
<thead>
<tr>
<th>Formula</th>
<th>Longitude of centroid (°E)</th>
<th>Latitude of centroid (°N)</th>
<th>Distance to the weighted centroid on raster image (m)</th>
<th>Range error to echo center (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula(2)</td>
<td>116.67054759</td>
<td>24.19727409</td>
<td>525.0632</td>
<td>0.7648</td>
</tr>
<tr>
<td>Formula(5)</td>
<td>116.67054662</td>
<td>24.19727487</td>
<td>524.9313</td>
<td>0.7646</td>
</tr>
<tr>
<td>Formula(7)</td>
<td>116.67054564</td>
<td>24.19727564</td>
<td>524.8005</td>
<td>0.7644</td>
</tr>
</tbody>
</table>

From Table 1, it’s known that the three centroids derived from three methods are very close and the distances among the three points do not exceed one meter. This states that there is no point in considering the weight of the distance to radar center in actual calculations. However, the accuracy of calculation will be correspondingly higher by way of considering the distance weight. In addition, although the echo center on vector map is closer to the echo center on raster map, a big error which is about 500m still exists. The authors analyze errors that possibly root in the following three aspects:

(a) Error caused by the transfer of axes. During the conversion from raw echo data to the raster data, the transfer of axes is frequently executed from the polar coordinate to the rectangular coordinate. The core of the conversion algorithm is that the echo value of the grid is arranged by the echo value of the nearest cell under the polar coordinate system. So this conversion may cause certain errors. By observing Figure 5 and Figure 6, it’s seen that the echoes from the two figures are similar, while they are different on some regions of the maps.

(b) Error caused by the curvature of the earth’s surface. During the inversion of raster echoes, the changes of the coordinates of latitude and longitude are considered as homogeneous. Actually, because of curvature of the earth’s surface, the coordinates of latitude and longitude vary with the distance heterogeneously.

(c) Error caused by the different accuracy of data access. During the process of generating raster and vector echo maps, a lot of calculations are executed, and then the calculations will certainly influence the accuracy of data. In addition, the process of calculating weighted centroid also needs a lot of calculations. Thereby, the differences of accuracy will affect the result of calculating the centroid to some extent.

6. Conclusion

By combining the GIS techniques, it is meaningful for the development of characteristic tracking algorithms to study the generation of the vector echo data and the localization of
weighted centroid of dBZ on it. This paper principally discusses the conversion from the base reflectivity products of CINRAD/SA weather radar to the vector echo data in shp format and proposes some localization methods of weighted centroid of dBZ on it. By comparing with the calculated values on the raster echo maps, it also tests these methods on the vector echo maps and analyses the calculated results and error sources which are mainly caused by the transfer of axes, the curvature of the earth's surface and the different accuracy of data access. In the end, it draws two conclusions that the echo's precision in vector format is much higher than that in raster format and it is more accurate to take the features area and distance to radar center as weights during the localization of weighted centroid of dBZ on vector echo maps. However, the authors don’t study the characteristic tracking algorithms based on this type of echoes. So it will have more significance for the radar warning and forecast to study an algorithm that can identify, track and forecast storms based on the vector echo data. These works will be the further research focus of authors.

References