Automated Cloud Patch Segmentation of FY-2C Image Using Artificial Neural Network and Seeded Region Growing Method (ANN-SRG)

By Yu Liu, Du-Gang Xi, Xue-Gong Liu, Chun-xiang Shi & Kai Zhang

Tianjin Hydraulic Research Institute

Abstract - This paper presents a new algorithm Artificial Neural Network and Seeded Region Growing (ANN-SRG) to segment cloud patches of different types. This method used Seeded Region Growing (SRG) as segmentation algorithm, and Artificial Neural Network (ANN) Cloud classification as preprocessing algorithm. It can be trained to respond favorably to cloud types of interest, and SRG method is no longer sensitive to the seeds selection and growing rule. To illustrate the performance of this technique, this paper applied it on Chinese first operational geostationary meteorological satellite FengYun-2C (FY-2C) in three infrared channels (IR1, 10.3-11.3μm; IR2, 11.5-12.5μm and WV 6.3-7.6μm) with 2864 samples collected by meteorologists in June, July, and August in 2007. The result shows that this method can distinguish and segment cloud patches of different types, and improves the traditional SRG algorithm by reducing the uncertainty of seeds extraction and regional growth.

Keywords : FY-2C; multi-channel satellite image; Artificial Neural Network; Seeded Region Growing; cloud patch; segmentation; cloud type.

GJCST Classification: I.2.1
Automated Cloud Patch Segmentation of FY-2C Image Using Artificial Neural Network and Seeded Region Growing Method (ANN-SRG)

Yu Liu *, Du-Gang Xi *, Xue-Gong Liu p, Chun-xiang Shi Q & Kai Zhang ¥

Abstract - This paper presents a new algorithm Artificial Neural Network and Seeded Region Growing (ANN-SRG) to segment cloud patches of different types. This method used Seeded Region Growing (SRG) as segmentation algorithm, and Artificial Neural Network (ANN) Cloud classification as pre-processing algorithm. It can be trained to respond favorably to cloud types of interest, and SRG method is no longer sensitive to the seeds selection and growing rule. To illustrate the performance of this technique, this paper applied it on Chinese first operational geostationary meteorological satellite FengYun-2C (FY-2C) in three infrared channels (IR1, 10.3-11.3μm; IR2, 11.5-12.5μm and WV 6.3-7.6μm) with 2864 samples collected by meteorologists in June, July, and August in 2007. The result shows that this method can distinguish and segment cloud patches of different types, and improves the traditional SRG algorithm by reducing the uncertainty of seeds extraction and regional growth.

Keywords . FY-2C, multi-channel satellite image, Artificial Neural Network, Seeded Region Growing, cloud patch, segmentation, cloud type.

I. INTRODUCTION

Clouds play an important role in the earth–atmosphere system. They significantly affect the heat budget by reflecting short-wave radiation (Hobbs 1981), and absorbing and emitting long-wave radiation (Hunt 1982). Cloud segmentation based on satellite images is vital for the parameter extraction of clouds patch and cloud track which are useful for numerous climatic, hydrologic and atmospheric applications. Therefore, accurate and automatic cloud segmentation has been a great interest of many scientists.

A wide variety of methods and algorithms are available to deal with the automatic segmentation of images (Fu and Mui 1981, Pal and Pal 1993). These methods can be broadly classified into four categories (Zhu and Yuille 1996): Edge-based techniques, (Region-based techniques, Deformable models, Global optimization approaches. Considering various shapes and blurry edges of cloud and the influence of underlying land surface, region-based techniques (Simpson et al. 1998, Shin et al. 1996) is suitable to detect clouds. For example, Seeded Region Growing algorithm (SRG) (Adams and Bischof 1994), exploit the spatial information of pixels and the formed regions will be homogeneous and connected. However, Conventional region-growing methods depend on the consistency of the image, and are sensitive to seed initialization and thresholds (Zucker 1976, Wei Hong et al. 2008). It is usually limited to regional model because cloud top bright temperature (TBB) varies greatly from different region and season.

Almost all those method mentioned before usually use only one infrared image. Plenty information provided by other channels are wasted. Most important of all, there is no method to identify cloud patches of different types directly. To deal with this problem, this study tries to use Seeded Region Growing (SRG) as segmentation algorithm and pre-process the satellite images by using pixel-based classifiers based on multi-channel data. Because cloud classifiers can reduce the number of satellite image value from hundreds floats to several integers. In addition, it is possible to determine cloud types according to research objective.

As for pixel-based classifiers, a lot methods have been developed for remote-sensing instruments using various machine learning techniques, such as neural network (Key et al.1989), Bayesian methods...
ANN-SRG technique we proposed is an iterative process by which regions are merged starting from individual pixels, or another initial segmentation, and growing iteratively until every pixel is processed. It can be described by the following steps:

1) **ANN Cloud Classification Model**

   In order to simplify image analysis and computer vision for the subsequent segmentation step, this study pre-processed satellite image by classifying cloud of FY-2C data with ANN in three infrared channels (IR1, IR2, WV). Seven categories have been divided: sea, land, low-level clouds, middlelevel clouds, thin cirrus, thick cirrus, multi-layer clouds, cumulonimbus. Its detail description of the ANN cloud classifier of FY2C can be seen in Yu Liu(2009). The schematic diagram of the result can be shown as Fig.2(b).

2) **Seeded Region Growing**

   The seeded region growing technique is an iterative process by which regions are merged starting from individual pixels, or another initial segmentation, and growing iteratively until every pixel is processed. It can be described by the following steps:

   1) **Pre-processing**: Segment the entire image into pattern cells. Details are described in the previous part: B(1). **ANN Cloud Classification Model**.

   2) **Seed extraction**: Scan image and find out a cell that hasn’t been labeled. This step try to find out seed (CTseed) whose cloud type is desired (CTdesired) and hasn’t been labeled. As it is shown in Fig.2(c), this method is not sensitive to seed positioning given a certain cloud type;

   3) **Region growing**: Each pattern cell is compared with its neighboring cells to determine if they are the same type (CTi=CTseed). If they are same, merge the cells to form a fragment and update the property used in the comparison. Then Continue growing the fragment by examine all of its neighbors until no joinable regions remain. Label the fragment as a completed region. The schematic diagram of the result of ANN-SRG (Fig.2(c)) shows that this method not sensitive to region growing principal.

   4) **Unlabelled Parts Selection**: Selects unlabelled parts to reapply steps 2), 3).
The four steps are iterated until the whole image is labeled.

Fig. 1: Configuration for the cloud segmentation; CTseed - Cloud type of seed; CTdesired - Desired cloud type; CT_{i,r} - Cloud type of pixel (i, j).

Fig. 2: Schematic diagram of the result of ANN cloud classification and ANN-SRG segmentation
(a) Original satellite map. Different color shows different cloud characteristics. (b) Result of ANN cloud classification. A, B, C and D are cloud types. (c) Result of ANN-SRG segmentation. A1 (A1, A2...An) is the cloud patch i whose cloud types is A. B1, C1 and D1 are the first cloud patch whose cloud types are B, C and D respectively.

(c) Evaluating Model Performance
The performance of the ANN-SRG can be judged by the three aspects: time cost, robustness, precision. The first one can be demonstrated by the training time consumed.

To analyze the robustness of the model, this study used different percentage (10%, 20%, 30%, 40%, and 50%) of validation cases, and performs the model under the existence of external disturbances (5%, 10%, 15% of fraud samples in training and validation phase). Because the uncertainty of the ANN-SRG method only comes from ANN classification as mentioned before, this study mainly evaluates the precision of ANN classifier. Its performance can be indicated by confusion matrix for the test data of cloud classifier, and some indexes, such as mean square error (MSE), normalized mean square error (NMSE), error (%), correlation coefficient. These indexes are defined as the following:

1) Mean square error (MSE)
\[
MSE = \frac{1}{N \times P} \sum_{i=1}^{N} \sum_{j=1}^{P} (d_{ij} - y_{ij})^2
\]

Where \(d_{ij}\) is desired output for exemplar \(i\) at processing elements \(j\); \(y_{ij}\) is network of output for exemplar \(i\) at processing elements \(j\); \(N\) is number of exemplars in the data set; \(P\) is number of output processing elements.

2) Normalized mean square error (NMSE):
\[
NMSE = \frac{P \times N \times MSE}{\sum_{i=1}^{N} \sum_{j=1}^{P} (d_{ij} - \overline{d}_i)^2}
\]
Where $d_{ij}$ is renormalized network of output for exemplar $i$ at processing elements $j$; $dd_{ij}$ is renormalized desired output for exemplar $i$ at processing elements $j$.

3) **Percent Error (%):**

$$\text{Error} = \frac{100 \times \sum_{j=0}^{N} \sum_{i=0}^{P} \left| \frac{dy_{ij} - dd_{ij}}{dd_{ij}} \right|}{N \times P}$$  \hspace{1cm} (3)

4) **Correlation coefficient (Corr):**

$$r = \frac{\sum_{j=0}^{N} \sum_{i=0}^{P} (d_{ij} - \bar{d})(y_{ij} - \bar{y})}{\sqrt{\sum_{j=0}^{N} \sum_{i=0}^{P} (d_{ij} - \bar{d})^2} \sqrt{\sum_{j=0}^{N} \sum_{i=0}^{P} (y_{ij} - \bar{y})^2}}$$  \hspace{1cm} (4)

Where $d$ is desired output for exemplar; $\bar{y}$ is network of output for exemplar.

### III. EXPERIMENTAL RESULTS

The segmentation result of 6 types of cloud patches (Low cloud, Medium cloud, Thin cirrus, Thick cirrus, Multi-layer cloud, Cumulonimbus) are presented in Fig.3.

**Fig.3:** (a) pseudo-color composite map of Tbb of IR1, IR2 and WV(16:00 UTC on 9 July 2007); (b) cloud classification results of ANN; (c) Segmentation result of low cloud; (d) Segmentation result of medium cloud; (e) Segmentation result of thin cirrus; (f) Segmentation result of thick cirrus; (g) Segmentation result of multi-layer cloud; (h) Segmentation result of cumulonimbus.

### a) ANN-SRG Segmentation Precision

The segmentation precision is analyzed from the pixel level and the image level which is given with the comparison of traditional SRG.

#### 1) Segmentation precision of pixel

The model errors is 8.87% in test period and the MSE, NMSE and correlation rates (Corr) of test and cross-examination of of ANN are 0.01, 0.02, 0.99
respectively. It shows that ANN-SRG methods can classify and segment clouds well.

The confusion matrix in Table 1 demonstrates that the model can differentiate 87%-98% of Low-level clouds, Midlevel clouds, Thin cirrus, Thick cirrus, Multi-layer clouds, Cumulonimbus accurately. The misjudgment always related to the similarity of clouds temperature, for example, it misjudged multi-layer clouds as cumulonimbus and thick cirrus which all of them has low temperature.

Table 1. Confusion matrix for the test data of cloud classifier

<table>
<thead>
<tr>
<th>Classes</th>
<th>Sea</th>
<th>Land</th>
<th>Low-level clouds</th>
<th>Midlevel clouds</th>
<th>Thin cirrus</th>
<th>Thick cirrus</th>
<th>Multi-layer clouds</th>
<th>Cumulonimbus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sea</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Land</td>
<td>0</td>
<td>0.97</td>
<td>0</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low-level clouds</td>
<td>0.04</td>
<td>0.05</td>
<td>0.87</td>
<td>0.02</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Midlevel clouds</td>
<td>0.01</td>
<td>0</td>
<td>0.02</td>
<td>0.92</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thin cirrus</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.93</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thick cirrus</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.02</td>
<td>0.92</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>Multi-layer clouds</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
<td>0</td>
<td>0.01</td>
<td>0.90</td>
<td>0.06</td>
<td>0</td>
</tr>
<tr>
<td>Cumulonimbus</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.01</td>
<td>0.98</td>
<td>0</td>
</tr>
</tbody>
</table>

2) Segmentation Precision of Image

Fig. 3 show that the configuration of cloud patches has smooth boundary, and low noise. They are close to the reality according to Meteorologists. The ANN-SRG method can detect different types of cloud well.

To compare the segmentation result of ANN-SRG with traditional SRG, this study chooses Mesoscale convective complexes (MCCs) case. Reasons for limit to the MCCs cases lie in the following aspects: firstly, MCCs are responsible for most of the warm-season rainfall (Augustine and Howard 1991); Secondly, According to the Physical mechanism of MCCs, it is composed of Cumulonimbus and Multi-layer clouds. The segmentation result by SRG and ANN-SRG can be compared; Thirdly, it is not possible to identify clouds with only one satellite channel data except MCCs whose TBB is less than -52°C according to its definition (Augustine and Howard 1991, Jirak et al. 2003). Therefore, this study use -52°C as the growth principle of traditional SRG.

Segmentation result of a Mesoscale convective complexes (MCCs) case of the traditional SRG and ANN-SRG can be seen in Fig. 4. It shows that ANN-SRG method can provide more detailed information by differentiating Cumulonimbus well from Multi-layer clouds. On the other hand, it improves the traditional SRG by reducing the uncertainty of seeds extraction and regional growth.
Fig. 4: Segmentation result of the SRG and ANN-SRG (16:00 UTC on 9 July 2007). (a) Infrared map of Channel 1; (b) Segmentation result of MCCs by SRG; (c) Segmentation result of Cumulonimbus by ANN-SRG; (d) Segmentation result of Multi-layer cloud by ANN-SRG

b) Robustness of ANN-SRG
The robustness of the ANN-SRG model for different validation samples and some fraud samples can be seen in Table 3. It clearly suggests that error of the model is about 10% in test period with 5-50% (Table 2 Model 1) validation samples and 5-15% (Table 3 Model 1) fraud samples for validation and the training.

Table 2: The result of cross-examination and tests of cloud classification model for different validation samples and for some fraud samples

<table>
<thead>
<tr>
<th>Percent</th>
<th>Cross-examination</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE</td>
<td>NMSE</td>
</tr>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>10%</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>20%</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>30%</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>50%</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%-5%</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>10%-5%</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>15%-5%</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>5%-10%</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>10%-10%</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>15%-10%</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>5%-15%</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>10%-15%</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>15%-15%</td>
<td>0.01</td>
<td>0.04</td>
</tr>
</tbody>
</table>

The classification result of ANN for different validation samples. The sample for all models in this table consists of 2864 cases. For the 5% model, the training sample is 95% cases (2721) and the validation sample is 5% cases (143). The others also correspond to this system so that the 15% model has 15% (429) cases in its validation sample.
The classification result of ANN with some fraud samples. The model trains on the 2292 cases and uses the last 572 as a validation. For the 5%-5% case, the training sample is 5% fraud cases (114) and 95% non-fraud cases (2178). The validation sample is 5% fraud cases (28) and 95% non-fraud cases (544). The others also correspond to this system so that the 5%-15% model has 5% fraud cases (114) in training sample and 15% (85) fraud cases in its validation sample.

c) Execution Time

We have measured the execution times of SRG and ANN-SRG (Table 3). It shows that the SRG consumes only 66 seconds to segment MCCs, while ANN-SRG consumes about 5 minutes for each cloud patches. However, the results of ANN-SRG are acceptable because FY-2C real time infrared images are collected hourly. Therefore, there is no obvious time limit for the proposed method.

The results also demonstrate that the pre-processing period of ANN-SRG occupies the most of time during segmentation process. Further improvement should be carried out. The execution time of different types of cloud is quite different. The fastest one is the segmentation of cumulonimbus, because of it its low frequency of occurrence on satellite image compared to other types of clouds.

<table>
<thead>
<tr>
<th>Type</th>
<th>Execution time</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>225</td>
</tr>
<tr>
<td>Low cloud</td>
<td>62</td>
</tr>
<tr>
<td>Medium cloud</td>
<td>50</td>
</tr>
<tr>
<td>Thin cirrus</td>
<td>53</td>
</tr>
<tr>
<td>Thick cirrus</td>
<td>54</td>
</tr>
<tr>
<td>Multi-layer cloud</td>
<td>52</td>
</tr>
<tr>
<td>Cumulonimbus</td>
<td>43</td>
</tr>
<tr>
<td>Total</td>
<td>539</td>
</tr>
<tr>
<td>SRG</td>
<td>MCCs</td>
</tr>
<tr>
<td></td>
<td>66</td>
</tr>
</tbody>
</table>

**Table 3.** Execution time FY-2C segmentation system(S)

**IV. Conclusion**

The study has presented a new segmentation scheme (ANN-SRG) to detect cloud patches of different type directly by the combination of artificial intelligent and traditional segmentation methods, and applied it to FY-2C in three infrared channels (IR1, 10.3-11.3μm; IR2, 11.5-12.5μm and WV 6.3-7.6μm). The results show that the ANN-SRG method can take advantage of multi-channel satellite data and segment cloud patches of different types which can’t be achieved by traditional segment method. It improves traditional SRG by reducing the uncertainty of seeds extraction and regional growth, because of using ANN cloud classifiers to reduce the number of infrared image value from hundreds floats to several integers.

The segment results of cloud patches obtained so far have been encouraging. Region boundaries have good correspondence with the contours of the cloud patches. However, in reality ANN-SSR method might not be able to achieve such good results as demonstrated in this study. There are mainly two reasons to constrain the further application of the current model. First, ANN-SSR method is sensitive to pre-processing result by ANN cloud classification which greatly depend on expert experience. Moreover, the infrared top brightness temperature (TBB) varies greatly in different region and seasons because of the influence of underlying land surface. To truly establish an effective and efficient cloud classification algorithm for the FY-2C satellite, more work needs to be carried out to analyze whether it is possible to build a cloud classification method independent of regions and seasons. Second, due to the relative complexity of ANN cloud classifier and series algorithm, the execution time of ANN-SSR is a little much, and its takes about 9 minutes for the cloud segmentation of each FY-2C image. Further improvement should be made, such as more quick cloud classifier and parallel algorithm.

**Acknowledgements**

Special thanks go to Pro.Xia,J.,and Ph.D Hong, Y.for their contributions to the manuscript.

**References**

Automated Cloud Patch Segmentation of FY-2C Image Using Artificial Neural Network and Seeded Region Growing Method (ANN-SRG)


