STRATEGIC ALTERNATIVES TO CALCULATE CROP AREA OF AN ADMINISTRATIVE DIVISION

Hongliang Fang

State Key Lab. of Resources and Environment Information System,
Institute of Geography, Chinese Academy of Sciences, 100101, Beijing, China

KEY WORDS: Agriculture, Estimation, Identification, Inventory

ABSTRACT:

In crop yield estimation using remotely sensed data, it usually needs to calculate the crop planting area in a particular administrative division. Most previous investigators do as follows: first, they cut down the target image of the study area with the administrative boundary, then conduct land cover/use classification and crop identification work, and last calculate the crop area. Other fewer researchers conduct the land cover/use classification work first and then cut down the study area with administrative boundary, and last calculate the crop area. We call these two methods strategy A(cut and classify) and strategy B (classify and cut) respectively. In this paper, we applied these two strategies to rice planting area identification. Our results indicate that strategy B is obviously excellent than strategy A in the unsupervised classification-cluster process and the rice area accuracy it extracted is over 84%.

1. INTRODUCTION

Crop production forecasting using remote sensing technique involves the estimation of the cultivated area and yield per unit area. Estimating the cultivated area is the primary step for crop yield forecasting(Tennakoon et al., 1992). In estimating the cultivated area, it usually needs to calculate the crop planting area of a particular administrative division for the purpose of being in agreement with statistical channels.

To calculate the crop planting area of a particular administrative division using remotely sensed data is computer and labor consuming. By using the map to determine the location of the administrative boundaries, we may be able to remove the inappropriate areas before beginning the bulk of image-processing computations, and thus decrease the costs of the analysis(Star and Estes, 1990). In fact, not every researcher does as above. Some perform image analysis first, then mask out the uninterested area with the key boundaries, and thus get the particular target information(Hall-konyves, 1990; Makey et al., 1993).

Thus, we have two strategies here: first, cut down the target image of the study area with the administrative boundary, and then conduct land cover/use classification and crop identification work, and last calculate the crop area. Second, conduct the land cover/use classification work firstly and then cut down the study area with administrative boundary, and last calculate the crop area of the study area. We call these two methods strategy A (cut and classify) and strategy B (classify and cut) respectively. Taking rice area identification as an example, the main aim of this paper is to compare these two strategies and to see which one is better and more robust.

2. PREVIOUS WORK

In this part, we retrieve some previous work on land cover/use classification and on crop area calculation in an administrative division. It showed that much of these work was done based on strategy A (cut and classify) using supervised or unsupervised automatic classification method as well as visual interpretation. Strategy B (classify and cut) was only used by fewer investigators, in whose cases supervised classification method was mainly applied.

2.1 Strategy A

In inland wetland change detection in the Everglades Water Conservation Area 2A(WCA-2A), Rutchey and Vilchek(1994) use Landsat Multispectral Scanner(Mss) data and SPOT High Resolution Visible(HRV) multispectral data to inventory aquatic macrophyte changes. Firstly, the polygon boundary of the study area was rasterized to a UTM map projection. The mask was applied to each of the rectified images. Thus only land within WCA-2A was allowed to contribute to the cluster development in the cluster development in the classification phase. A standard statistical unsupervised classification of the study area was performed, yielding 30 clusters and then aggregated to 7 classes(Jensen et al. 1995). Multitemporal classification results were used to identify the change in the spatial distribution of aquatic macrophyte.

Bauer et al.(1994) described the procedure to inventory Minnesota forest resources using multitemporal Landsat TM data. Because of the large study area, three scenes were mosaiced and then cut down the target area(5 counties). Consequently, unsupervised classification was conducted and 11 forest cover type map and their corresponding areas were got. Change trend analysiscan be conducted with multi-year areas. It should be noted that the last forest cover type map was not obtained from classifying the mosaiced TM data. Three separate images were processed individually and the results were merged as classified(GIS) files.

Strategy A was applied by previous researchers not only in unsupervised classification but also in supervised classification. Ray et al.(1994) estimated cotton production in India using IRS-1B and meteorological data. After mosaicing two scenes of LISS -1 data, study area boundary pixels were used to cut down the target image. Maximum likelihood classification scheme was used for crop identification and acreage estimation. The same strategy was applied in Turkey in acreage estimation of wheat and barley(Pestemalei et al., 1995).

2.2 Strategy B

Contrary to the above most authors who use strategy A in their experiment, fewer authors applied strategy B in their project. Hall-Konyves(1990) introduced basic principals and techniques in Sweden to monitor the crop. Maximum classification technique was applied by the author and to improve the accuracy of the classifier, a post-classification clean-up filter was applied. The author states "no area was masked out prior to classification".

In the pro-harvest state level wheat acreage estimation of Punjab, India, Makey et al.(1993) obtained the wheat acreage for six stratausing supervised classification. But the author hasn't mentioned which strategy was applied.

3. EXPERIMENTAL

Since strategy A has been applied by most pervious researchers insupervised and unsupervised classification and strategy B in supervised classification, the main aim of our experiment is to see the effect of strategy B in unsupervised classification.

3.1 Study Area and Data

The study area we selected is the county of Jiangling in Hubei province, China. Jiangling county, located in middle Yangtze River Plain, is a major rice production county in Hubei province. The early rice is sowed in the third decade of March or the first decade of April and transplanted in the third decade of April of the first decade of May. The semilate rice is sowed in the first decade of May and transplanted in the third decade of June. According to the farming practice of the study area, Landsat-5 TM CCT, dated 8 June, 1992 when it was clear and cloudless, scene of Path-124, Row-39 containing the whole county, was acquired from the Chinese Satellite Ground Station. The image processing system we used is ERDAS software and ARC/INFO GIS software is also supplementally used. Moreover, 1:50 000 scale topographic maps, recent vegetation type maps, soil maps, land cover/use maps and other ancillary information were available.

3.2 Experimental Procedure

The whole experimental process includes 5 steps:

Step 1 : load Landsat-TM digital data(scene 124-39) to Sun workstation(INPUT). Ground Control Points(GCPs), extracted from 1:50 000 topographic map in Gauss-Kreuger projection system, were used to conduct geometrical correction. Radiometric correction hasn't been done.

Step 2 : display the administrative boundary of Jiangling county(ARC/INFO vector format) on the screen and cut down(CUTTER) the rectangular image circumscribing the study area with a rectangle box. The administrative boundary of Jiangling county was extracted form the same topographic map as did GCPs. Plate 1 is a standard false-color composite image of the rectangular image containing the whole study area. The yellow line is the administrative boundary of Jiangling county.

Step 3 : for strategy A, mask out the inappropriate area outside the administrative boundary. A standard statistical unsupervised classification was performed(CLUSTR) and the target area was clustered into 50 clusters.

Step 4 : for strategy B, unsupervised classification was conducted(CLUSTR) directly and the rectangular area was clustered into 50 clusters also. Then, cut out the inappropriate area with the administrative boundary.

Step 5 : referring to the soil maps, land cover/use maps and topographic maps, the unsupervised classification result from step 4 was recoded into 10 major land cover types.

3.3 Results and Analysis

Table 1 and table 2 list the unsupervised classification result of strategy A and strategy B respectively. Plate 2 illustrates the condensed 10 clusters from step 5. In plate 2, red color stands for early rice and pink semilate rice, shallow gray irrigated land and blue water.

In table 2 only 14 clusters occupied more than 2 percent of the total non-zero pixels, while in table 1 there are 17 clusters. This means pixels within the administrative boundary from strategy B has a better accumulation effect compared with that from strategy A which ismore evenly distributed. In table 1 of strategy A, the first 6 clusters occupied more than 43 percent of the total non-zero pixels. It is hard to avoid class confusion when most pixels are aggregated inthe first several clusters.In table 2 of strategy B, the first 6 clusters only occupied 23.28 percent of non-zero pixels. The peaks of the histogram are secretely distributed at cluster #7, #9#19, #22, #25, #32, #47 respectively. This means the main land cover types have not aggregated in the first several clusters(table 2) but separately distributed. Having identified the corresponding land cover type of these peaks, it is convenient to evaluate and label surrounding 'non-peak' clusters. The extracted rice area accuracy with strategy B is over 84% compared with the statistics obtained from the Agricultural Investigation Team affiliated to the Agricultural Bureau of Hubei province.

4. DISCUSSION

In supervised classification, the spectral characteristics of the training sites are used to "train" the classification algorithm for eventual land cover mapping of the remainder of the image(Jensen, 1986). Once representative training sites selected, multivariate statistical parameters calculated from each training site are used to evaluate every pixel. The pixel is assigned to the class of which it has the highest likelihood of being a member no matter it is within or outside the administrative boundary. As far as strategy A and strategy B are concerned, there is not much difference between these two strategies for supervised classification.

In a unsupervised classification, the computer is allowed to select the class criteria such as means and covariance matrices(Jensen, 1986). In strategy A, class criteria are calculated within the administrative boundary only while it is done within the whole circumscribing rectangle in strategy B. Certainly, diverse results will emerge for unsupervised classification with different strategy.

In image data, the spatial dependence among pixels is the fundamental aspect of spatial pattern(Henebry, 1993). If the satellite image was masked with boundary pixels, information about this dependence is lost at all(Henebry, 1993). According to this rationale, contrary to the method used by many authors, strategy B is used in our project. Statistical unsupervised classification work was performed firstly and then mask out the uninterested area outside the administrative boundary. By this way, we hope, the spatial dependence can be retained especially when it is concerned with the rice theme. Our results suggest that strategy B is practical in rice identification and obviously excellent than strategy A for unsupervised classification.

In the above discussion, no matter which strategy was used, the deviation induced by the boundary pixels should be noticed. Readers can refer to Crapper(1984) on how to calculate the boundary pixels. Rao and Mohankumar(1994) explained in detail the effect of spatial resolution and the percentage of boundary pixels on accuracy of area estimation. In general usage, boundary pixels when occupying only a little percentage of the total study area(0.442%, in our experiment), can usually be ignored in acreage estimation.

For visual interpretation, uncertainty in assigning a theme to a grid cell can occur when the grid cell lies near or on the boundary of a region(Crapper, 1984). According to this rationale, strategy B should also be considered in visual interpretation. That's to say, first, conduct the visual interpretation work and then mask out the uninterested area and calculate the labeled results.

5. CONCLUSIONS AND FURTHER RESEARCH

For strategy A, we only give the result of pre-classified 50 clusters. Further labeling and interpretation work haven't been done. Surely, it can be done, but certainly, it will be more complex than strategy B and the accuracy won't be satisfying.

In our test, the study area was also classified into 70 clusters for strategy A and strategy B in unsupervised classification. Similarly, with strategy B, the formerly classified 70 clusters can be more easily reclustered into 10 types than with strategy A.

In summary, it may be said that strategy B is far more suitable and robust in the unsupervised classification-recoding process than is strategy A. Strategy B, rather than strategy A, should be of first consideration in similar projects. Applying this strategy in our study, the rice area accuracy we identified is exceeding 84%.

We tested the effect of Landsat Thematic Mapper(TM) data in our experiment. For other remotely sensed data, such as NOAA-AVHRR image or SPOT image, further experiment is still needed. It is expected that similar results should be obtained and strategy B be more robust in these instances.

In practical usage, the rice planting area is always much lager than that in our case. A larger crop cultivated area may cover several TM scenes which have different ground spectral and spatial characteristics. Although strategy B is properly used in one scene circumstance, problems will occur if it is applied to multi-scenes. How to cope with this problem still needs further research.

REFERENCES

Bauer,M.E.,Burk, T.E.,Ek, A.R., Coppin, P.R.,Lime, S.D., Walsh, T.A., Walters, D.K., Befort, W., and Heinzen, D.F., 1994. Satellite Inventory of Minnesota Forest Resources. Photogramatric Engeering and Remote Sensing, 60(3), pp.287-298.

Crapper,P.F., 1984. An Estimate of the Number of Boundary Cells in a Mapped Landscape Coded to Grid Cells. Photogramatric Engeering and Remote Sensing, 50(10), pp.1497-1503.

Hall-Konyves,K., 1990. Crop monitoring in Sweden. International Journal of Remote Sensing, 11(3), pp.461-484.

Henebry, G.M., 1993. Detecting Change in Grasslands Using Measures of Spatial Dependence with Landsat TM Data. Remote Sensing of Environment, 46(2), pp. 223-234.

Jensen, J.R., 1986. Introductory Digital Image Processing. Englewood Cliffs, New Jersey, pp. 177-230.

Jensen, J.R.,Rutchey, K.,Koch, M.S.,and Narumalani, S., 1995. Inland wetland change detection in the Everglandes Water Conservation Area 2A using a time series of normalized remotely sensed data. Photogramatric Engeering and Remote Sensing. 61(2), pp.199-209.

Makey, P.K., Singh, R., Sidhu, S.S., Narang, R.S., Dadhwal, V.K.,Parihar, J.S.,and Sharama, A.K.,1993. Pre-harvest state level wheat acreage estimation using IRS-1A LISS-1 data in Punjab(India). International Journal of Remote Sensing, 14(6), pp.1099-1106.

Pestemalei,V., Dinc, U., Yegingil, I., Kandirmaz, M., Cullu, M.A., Ozturk, N., and Aksoy, E., 1995. Acreage estimation of wheat and barly fields in the province of Anana, Turkey. International Journal of Remote Sensing, 16(6), pp.1075-1085.

Rao, P.P.N. and MohanKumar,A., 1994.Cropland Inventory in the Command Area of Krishnarajasagar Project Using Satellite Data. International Journal of Remote Sensing, 15(6), pp.1295-1305.

Ray,S.S., Pokharna, S.S., and Ajai, 1994. Cotton Production Estimation Using IRS-1B and Meterological Data. International Journal of Remote Sensing, 15(5), pp.1085-1090.

Rutchey,K., and L.Vilchek, 1994. Development of an Everglades Vegetation Map Using a SPOT Image and the Global Positioning System. Photogramatric Engeering and Romote Sensing, 60(6), pp.767-775.

Star J. and John Estes, 1990. Geographic Information Syatem:An introduction. Prentice Hall, Englewood

Cliffs, New Jersey, 218 p.

Tennakoon, S.B., and Murty, V. V. N., and Eiumnoh, A., 1992. Estimation of cropped area and grain yield of rice using remote sensing data. International Journal of Remote Sensing, 13(3), pp. 427-439.