A KNOWLEDGE-BASED APPROACH FOR REDUCING CLOUD AND SHADOW

Mingjun Song, Graduate Research Assistant
Daniel L. Civco, Director
Laboratory for Earth Resources Information Systems
Department of Natural Resources Management & Engineering
The University of Connecticut
U-87, 1376 Storrs Road
Storrs, CT 06269-4087
mjsong@canr.uconn.edu
dcivco@canr.uconn.edu

ABSTRACT

Completely cloud-free remotely sensed images are not always available, especially in tropical, neo-tropical, or humid climates, posing complications and perhaps serious constraints to image analysis. In this paper, a knowledge-based approach to reduce cloud and shadow using two dates of Landsat TM imagery is presented. Procedurally, first, the brightness and contrast of a secondary image was adjusted to be the same as the main image. Second, instead of detecting clouds and shadows in two images independently, a knowledge base was used to detect their presence in the main image in areas not present in the secondary image. Corresponding brightness value thresholds for cloud and shadow area in the main image and non-cloud and non-shadow area in the secondary image were set in TM bands 1 and 4, respectively. Thresholds of brightness difference between two images were used to avoid other features having similar brightness values. An object-based method was employed to distinguish linear water features from shadow areas. *eCognition* was used to segment images into objects, with emphasis on water and shadow features. and length-to-width ratio of these objects was derived to separate confused linear water and shadow area. Finally, a composite image was generated with minimal cloud and shadow by replacing the brightness values of detected areas in the main image with those of the secondary image. This paper describes the method and presents results for a study area in Madagascar.

INTRODUCTION

Cloud-free remotely sensed images acquired from earth orbiting satellites are not always available, especially for areas of the earth characterized by tropical or humid climates. The problem of obtaining relatively-cloud free image mosaics is compounded for nadir-only observing satellites, such as Landsat, with a relatively infrequent revisit period. One such area in which these investigators have been conducting research on deforestation is the eastern coast of Madagascar where even a single scene (path-row) of TM or ETM data free from clouds and shadows is rare. However, procedures by which the effects of clouds and shadows can be reduced to aid in generating a cloud-free or relatively cloud-free remote sensing image is rarely studied. In this project, we adapted aspects of expert systems used for land use and land cover classification, to develop a knowledge-based method to produce a cloud and cloud shadow-free multitemporal image composite of neo-contemporary images (*e.g.*, within two to three months of one another).

DATA

The study area is located on the eastern coast of Madagascar where research is being conducted by one of the authors and his colleagues. The investigators are characterizing the interactions between the standing forests along the eastern coast of Madagascar and the contribution of human populations to their conservation, transformation and degradation. Physical, natural, sociopolitical, economical, demographical and historical information are being gathered and integrated using a Geographical Information Systems to test hypothesized factors affecting the changes in forests. The project is a collaboration between the University of Connecticut (Department of Anthropology, Department of Ecology and Evolutionary Biology, Department of Natural Resource Management & Engineering) and the Université d'Antananarivo in Madagascar (Forestry Department of the School of Agronomy, and the University Museum).

Two Landsat TM images were collected on April 22, 2001 (Figure 1a) and December 15, 2000 (Figure 1b) respectively, both of which were covered with some quantity of cloud and shadow, most of it not overlapping. The geographic extent of the two images is from 49°11′39″ E to 49° 28′10″ E longitude, and from 17°12′00″ S to 17°36′01″ S latitude. Both were projected into Laborde Oblique Mercator. A digitized contour map with an interval of 25 meters was used to generate a digital elevation model (DEM). As will be illustrated, these data were useful in assisting in producing a relatively cloud-shadow free composite image from these two Landsat images.

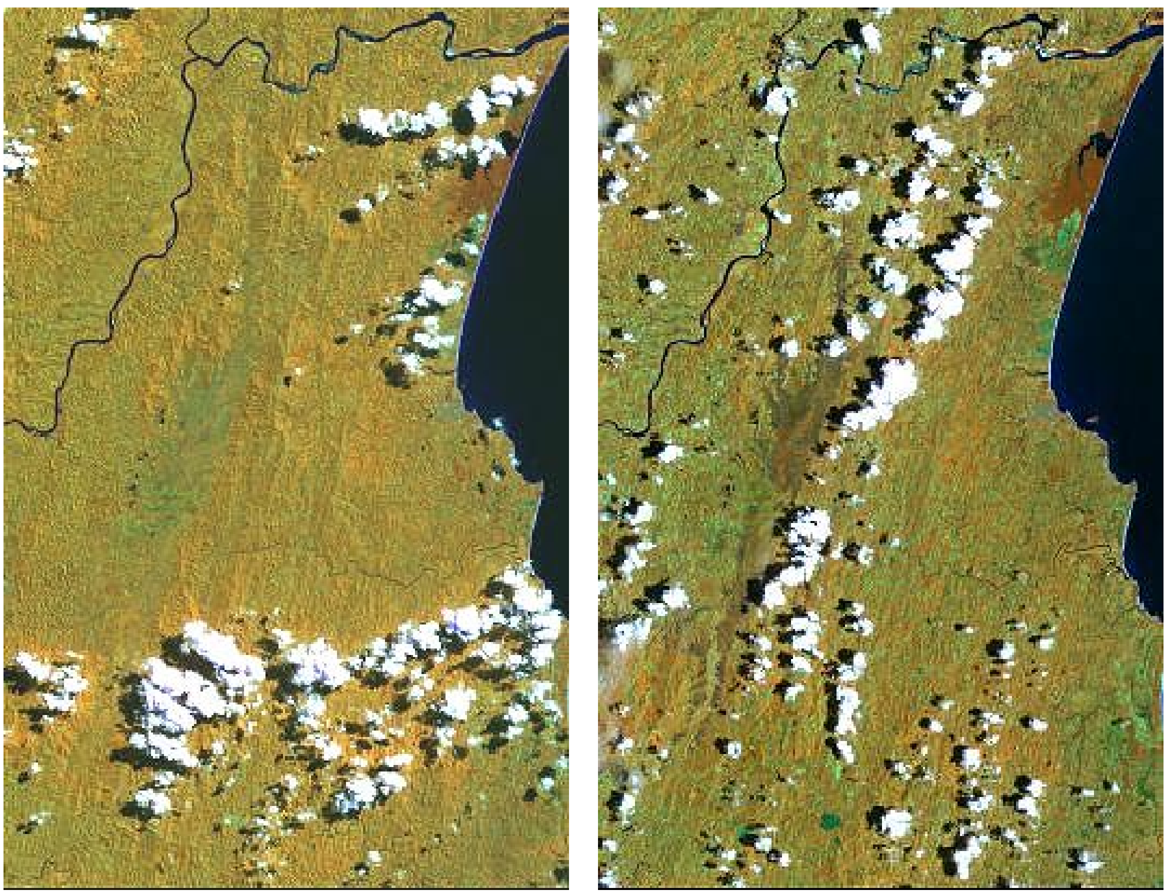

Figure 1. Landsat TM images of the eastern Madagascar study area. Bands 4, 5, and 3 are portrayed as R-G—B, respectively. 1a. (left) April 22, 2001. 1b. (right) December 15, 2000.

METHODS

In this research, topographical normalization was first performed to remove the effect s on spectral reflectance introduced by differential illumination and shadowing caused by terrain variability. Secondly, corrections were applied to account for brightness differences caused by different atmospheric conditions, moisture, and sun angles between the two dates. Then, a knowledge based approach was developed to reduce the cloud and shadow using two images of different dates. In this paper, the *main image* refers to the principal image to be used in subsequent analysis, such as land cover classification, and which possesses less cloud and shadow area than the *secondary image*, used to supplement the values for cloud and shadow areas the main image. The idea is to detect those areas which are covered with cloud and shadow in the main image, and not having cloud or shadow in the secondary image. This is accomplished using a combination of spectral information and shape features. Finally, the spectral values of cloud and shadow areas in the main image were replaced with those from cloud-free and shadow-free areas in the secondary image.

Topographical Normalization

Because of the topographic effect, similar land features at different slopes and aspects often have different spectral reflectance. To reduce this effect, the method developed by Civco (1989) was employed here. The basis to the algorithm is to adjust the spectral values according to the cosine of the incident solar angle. This process can be formulated as:

$$DN_{norm} = DN_{orig} + DN_{orig} \times (1 - \gamma / \gamma_{mean})$$

where DN_{norm} is the brightness value after normalization, DN_{orig} represents original brightness value, γ is the relief value and γ_{mean} is the mean value for the whole relief image. The result was stretched to the range of [0, 255].

Multi-date Effect Brightness Correction

Because of the time difference, and assuming no significant phonological difference, the brightness of the secondary image should be adjusted to the same level as main image. Caselles (1989) developed a simple approach using the statistical characteristics of multi-date images, which can be expressed as:

$$DN_{corr} = mean_{main} + (DN_{\sec d} - mean_{\sec d}) \cdot SD_{main} / SD_{\sec d}$$

where DN_{corr} is the corrected brightness of the secondary image, $DN_{\sec d}$ is the original brightness from the secondary image, $mean_{\sec d}$ and $SD_{\sec d}$ are the mean and standard deviation of the secondary image, and $mean_{main}$ and SD_{main} are the mean and standard deviation of the main image. However, when using this method in the multi-date images with substantial amounts of cloud and shadow present, the statistical parameters of image features can be severely affected by their presence. So in this research, the location of cloud and shadow areas were first approximated in both dates of imagery and masked out in the calculation of statistical measures.

Cloud and Shadow Reduction

A knowledge-based method was developed to detect the cloud and shadow areas that existed in the main image while not in the secondary image, instead of identifying these features independently in the two images. Through observation of the spectral reflectance characteristics of objects, TM Band 1 was deemed best for detecting cloud and Band 4 was best for detecting shadow.

The Band 1 brightness threshold for clouds in the main image was found to be greater than 41, and brightness of non-cloud areas in the secondary image was less than 33. To avoid other features having similar brightness values to those of clouds, a tolerance threshold of 10 was used for the difference between the two dates of imagery. The rationale is that if the difference is less than a threshold, it should be the same object in two dates images. The brightness value of Band 4 for shadow in the main image was found to be less than 35, and non-shade-water area was greater than 27. Again, a threshold of 10 was used for the difference between two dates images.

It was found because of the mis-registration and the effect of clouds, water features, principally a river, were potentially confused with shadow even with a difference threshold. In order to remove this confusion, shape information was used to distinguish between them.

Shadow and water areas that have brightness values smaller than 35 in Band 4 of the main image were extracted. The image was imported into eCognition[1] software, segmented into objects, which were exported along with their length/width ratio attributes. In ArcView[2], the length/width attribute for each object was converted to a grid file, and finally imported into ERDAS[3] Imagine. It was found that the length/width shape metric for the river had a value greater than 9. The criteria used for discriminating cloud and shadow areas is summarized in Table 1.

Table 1. Knowledge Base for Detecting Cloud and Shadow Areas

Parameter	**Cloud**	**Shadow**
Output Value	2	1
Band 1 (Main Image)	> 41	> 0
Band 4 (Main Image)	> 0	< 35
Band 1 (Secondary Image)	< 33	< 33
Band 4 (Secondary Image)	> 0	> 27
Band 1 Difference	> 10	> 0
Band 4 Difference	> 0	> 10
Length-to-Width Ratio	> 0	> 9

Finally, the pixel brightness values for areas identified as cloud and shade area in the main image were replaced with the values from the secondary image. This result is shown in Figure 3.The flowchart of this process is drawn as Figure 4 in an ERDAS Imagine Spatial Modeler format.

It should be noted that not all clouds and shadows were removed given just these two dates of imagery, since some coincidental areas of the scenes were obscured by cloud and shadow presence. This *noise* possibly could be removed using yet another date of Landsat TM imagery. Nonetheless, image quality and interpretability has been improved substantially, especially in the coastal forests of interest to these investigators.

[1] Definiens Imaging GmbH, Trappentreustr. 1, 80339 München, Germany
[2] ESRI, 380 New York Street, Redlands, CA 92373-8100, USA
[3] ERDAS, 2801 Buford Highway, N.E., Atlanta, GA 30329-2137 USA

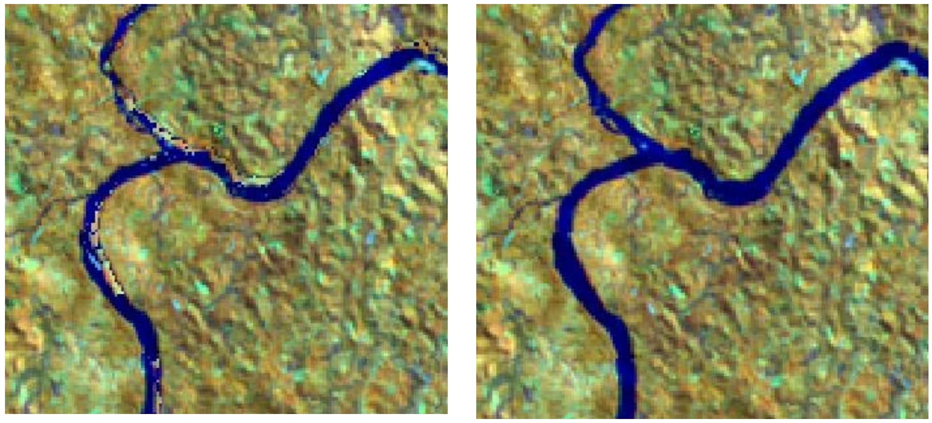

Figure 2. Enlargement of a portion of the composite image. Left: the result without shape information. Right: the result using the shape information.

Figure 3. The composite after cloud and shadow reduction

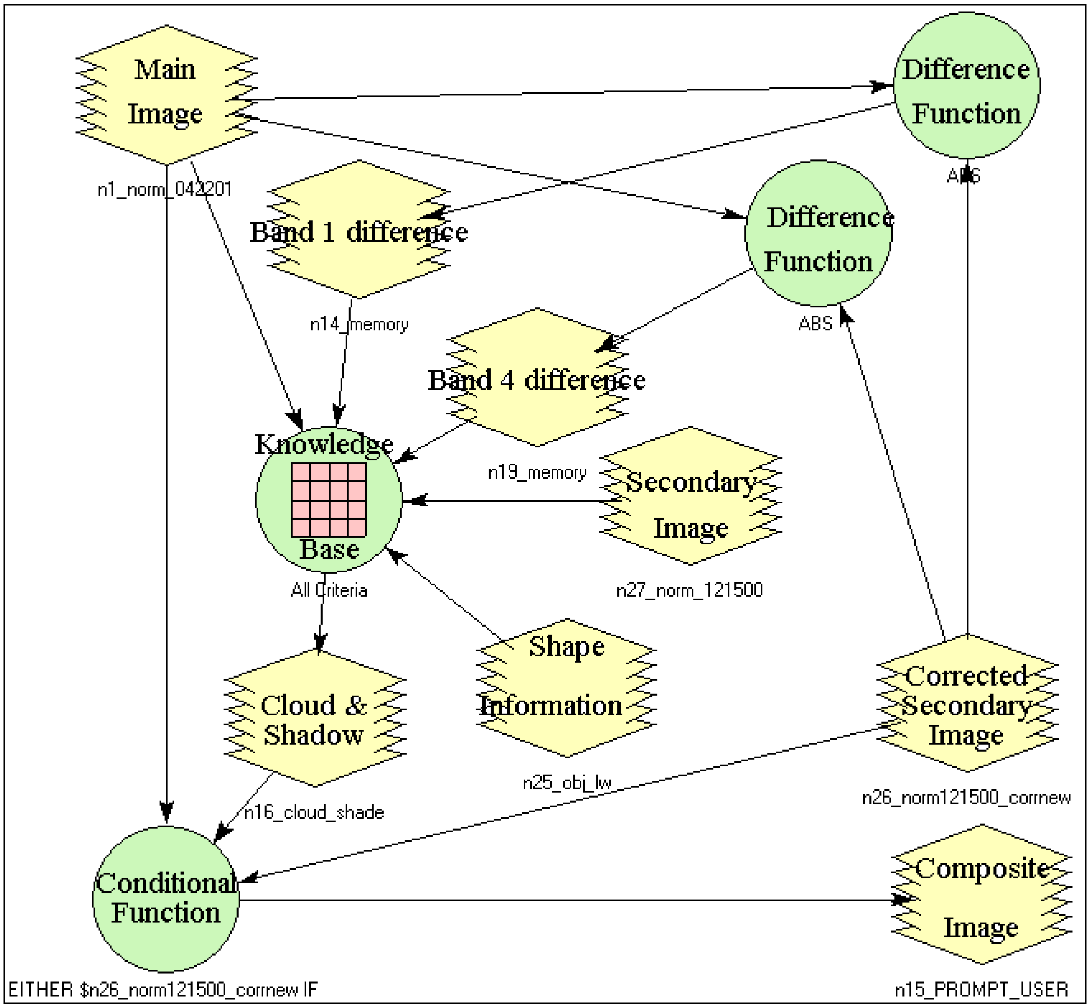

Figure 4. Flowchart of cloud and shadow reduction in Imagine Spatial Modeler format

CONCLUSION

In this research, a procedure of how to reduce the cloud and shadow effects to produce a mosaic of remotely sensed image with few clouds and shadows using neo-contemporary images was introduced, and a knowledge based approach to detect the cloud and shadow areas was developed. The experiment on the eastern coast of Madagascar shows that it is a simple and efficient method. The resulting composite has fewer cloud and shadow areas than either of the two dates of Landsat data alone. Another image of different date may be used again to remove the cloud and shadow still existing the composite. Furthermore, this method is flexible in detecting the cloud and shadow. For example, if the supplementary image is not severely cloud covered, the threshold for the band values of cloud areas may be reduced so that cloud border areas (edges) with less cloud influence can be replaced by values of non-cloud areas in the supplementary image. It is suggested that additional research be conducted to employ other *contextual* features in the expert system to improve the quality of the final product.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Aeronautics and Space Administration under Grant NAG13-99001/NRA-98-OES-08 RESAC-NAUTILUS, *"Better Land Use Planning for the Urbanizing Northeast:Creating a Network of Value-Added Geospatial Information, Tools, and Education for Land Use Decision Maker*s". CLEAR Publication Number 020115.4. SAES Scientific Contribution Number 2080.

LITERATURE CITED

Caselles, V. (1989). An alternative simple approach to estimate atmospheric correction in multitemporal studies. *International Journal of Remote Sensing*, 10 (6): 1127-1134.

Civco, D. L. (1989). Topographic normalization of landsat thematic mapper digital imagery. *Photogrammetric Engineering and Remote Sensing*, 55 (9): 1303-1309.