

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Cloud Context-based Onboard Data Compression

Y. P. Zhou

Science Systems and Applications, Inc.
Lanham, MD 20776, USA

P.-S. Yeh, W. J. Wiscombe, S.-C. Tsay

NASA Goddard Space Flight Center
Greenbelt, MD 20771, USA

Abstract- In this study, a simplified cloud mask algorithm suitable for present-day onboard processing is developed based on the operational MODIS cloud mask algorithm. The code is reduced to less than 45% of its original size while retaining 90% of the accuracy. Clear sky data compression is tested using our simplified cloud mask followed by application of the CCSDS lossless compression algorithm. For most of the channels, compressing the cloud masked MODIS L1B data can reduce the data volume by about 40% compared with compression without using cloud mask. The cloud mask was also applied to L1A data with comparable or even better data reduction performance depending on local data filling schemes. The cloud mask itself can be compressed to less than 0.5 bit-per-pixel. This paper summarizes our approach and provides results in detail.

Keywords- onboard data compression; MODIS; cloud mask

1. INTRODUCTION

Advanced land and ocean remote sensing platforms launched in the past decade have provided an enormous amount of information for scientists in their quest for understanding the Earth. However, in the visible and infrared range of the electromagnetic spectrum, the sensors also detect the signature of clouds which occur, on average, over fifty percent of the time. For most land and ocean studies, cloud-covered pixels are prevented from further processing into high-level products (such as aerosol, land-cover and sea surface temperatures). Nevertheless, these cloud-covered data must be transmitted to the ground and thus they impose unnecessary requirements on an already constrained communications channel.

As sensor technology advances towards ever increasing resolution (both spatially and spectrally), the data flow from space to ground will become increasingly clogged. Various onboard data processing techniques for data reduction have been proposed to alleviate the situation. Those implemented have been mainly signal-based techniques that take no account of the physical nature of the data. In this investigation, we focus on using the macro-scale context, the physical "cloud" information, for increasing onboard data reduction. This is because Earth observing data is over 50% cloud covered. The theoretical basis of the algorithm is well understood and is quite mature. Moreover, "cloud detection"

is the basis for various other potential onboard processing applications involving hyper-spectral sensors. In this investigation, we have constructed a simplified cloud mask algorithm based on Moderate Resolution Imaging Spectroradiometer (MODIS) operational cloud mask algorithm and demonstrated its application in data compression of clear sky measurements. Section 2 gives a brief introduction of MODIS cloud mask algorithm and the steps taken to simplify the MODIS cloud mask algorithm and its accuracy. Section 3 provides data compression results using cloud mask on MODIS level 1B and level 1A data. Section 4 summarizes the current work and future work in this direction.

II. THE SIMPLIFIED MODIS CLOUD MASK ALGORITHM

Clouds are generally characterized by higher reflectance and lower temperature than the underlying earth surface, so that simple visible and infrared window threshold approaches offer considerable skill in cloud detection. Difficulties arise when there is insufficient contrast between the radiance from cloud and from the underlying surface or surrounding atmosphere such as snow and ice surface, low stratus at night, or thin cirrus cases. Cloud edges present another difficulty since the instrument field of view may not see completely clear or cloudy scene. The cloud mask algorithms have been developed for the International Satellite Cloud Climatology Project (ISCCP) [1], the Advanced Very High Resolution Radiometer (AVHRR) under APOLLO (AVHRR Processing scheme Over Land, Cloud and Ocean) scheme [2, 3, 4, 5] and CLAVR (NOAA Cloud Advanced Very High Resolution Radiometer) algorithm [6]. The MODIS project adopted many of the existing cloud detection techniques and developed new spectral approaches in order to detect cloud in various situations [7]. The algorithm employs a series of visible and infrared threshold and consistency tests to specify confidence that an unobstructed view of the Earth's surface is observed.

MODIS has 36 spectral bands with center wavelengths ranging from 0.412 μm to 14.235 μm ; band 1-2 are imaged at a nominal resolution of 250m at nadir, band 3-7 are imaged at 500m, and the remaining bands at 1000m. Bands 13 and 14 each have two gain settings, low and high, telemetered from the instrument.

The current MODIS cloud mask algorithm utilizes 19 surface scene types and many of them are rather similar. To simplify the problem, we reduced the scene types from 19 to 7. The basic scene separations are between day and night, land and water. For daytime land, we treat desert and snow as separate scene types since these two are most unique (Fig.1). Without ancillary snow/ice input, a rough daytime snow algorithm by Hall et al. [8] is implemented in order to determine the snow scene for daytime, but nighttime snow is unavailable at this time. The simplified algorithm flow chart is defined in “scene types” and given in Fig. 1

MODIS uses multi-channel approach to determine the cloudy pixel from clear pixel. Five groups of tests are implemented in current operational algorithm. Each scene type uses a subset of these tests since not all of them are applicable to every scene type. Additional tests are applied when ambiguities are likely in the desert, coast, shallow-water, sun glint area. To decide tests to be used in the simplified algorithm, initial statistics have been performed on 477 granules (each granule contains an approximately 1034x2030 pixels of the size of 1 km²) to determine the performance of each test. These 477 granules include all the granules (288) on Day 222 of 2001 and selected granules from day 026 and day 187 of 2002 so that they provide more than a million pixels for each of the scene types and are well representative of different clouds and scene types. Several versions of algorithm are created and performance extracted and compared with results from the “operational” algorithm. Finally, a consolidated algorithm is generated. The consolidated algorithm uses a single test for four of the scenes (daytime snow, land, water and night time ocean), and two tests for nighttime land and Antarctic daytime land (Fig.1). The consolidated algorithm is less than half the size of the original algorithm. The result of the consolidated algorithm matches the operational cloud mask up to 90% for all the cases for both cloudy sky and clear sky as shown in Fig. 2. It should be noted that in extracting the cloud mask, pixels that are in the categories of *probable cloud*, *clear* or *probable clear* are marked as “non-cloud”. Thus a cloud mask can be represented by only two values: cloud or non-cloud, which requires only 1-bit representation.

III. DATA COMPRESSION WITH CLOUD MASK

Remote sensing of many surface properties only uses clear sky radiances, for these applications it is desirable to send only clear sky data. One way to reduce the clear sky data volume is to extract cloud mask and compress the masked data. The cloud mask itself can be compressed and transported as well. For compression, we use the CCSDS (Consultative Committee on Space Data System) lossless compression recommendation (extended-Rice algorithm) [9, 10] on MODIS Level 1B and Level 1A data.

A. L1B data

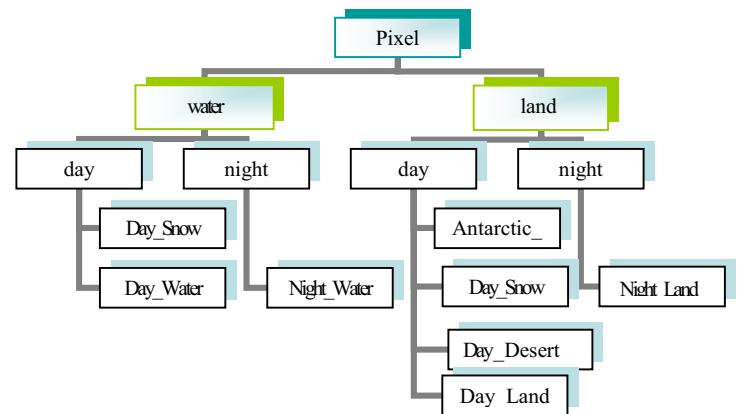


Figure 1. Flow chart with scene types.

The Level 1B data contains calibrated radiance/reflectance data in unsigned 16-bit format on which the cloud mask is generated [11]. A typical scene and the extracted cloud mask are shown in Fig. 3. The simplified cloud detection algorithm is applied on all 288 granules from day 222, 2001. This guarantees that the results are representative as the compression ratio (defined as 16/coded_bits_per_pixel) in Fig. 4 varies significantly with each image. For bands not saturated with clouds, the compression ratios are fairly stable (around 1.5 for visible, SWNIR channels and 1.7 for IR channels) without applying cloud mask. With cloud mask, compression ratios are improved 25-60% depending on bands and cloud amount in the image.

B. L1A data

L1A data [12] provide uncalibrated raw digital counts. The instrument produces 12-bit data. Thirty-seven granules are carefully selected to cover different surface types and cloud conditions. Applying the cloud mask extracted from L1B data to the L1A data provides realistic compression performance given in Figure 5 (compression ratio for L1A data is defined as 12/coded_bits_per_pixel). This data set also contains various amount of cloud coverage, with an average of 61%, typical of average MODIS daily coverage. The compression gain varies as the amount of cloud coverage changes as shown in Fig. 6.

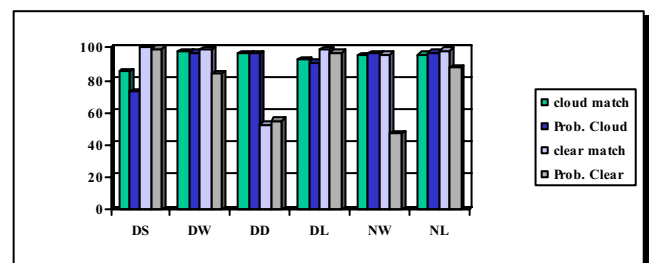


Figure 2. Performance of consolidated algorithm in percentage of correct matching pixels in *cloud*, *probable cloud*, *clear* and *probable clear* for DS(day_snow), DW(day_water), DD(day_desert), DL(day_land), NW(night_water), NL (night_land) scenes, based on 288 granules of day 222, 2001.

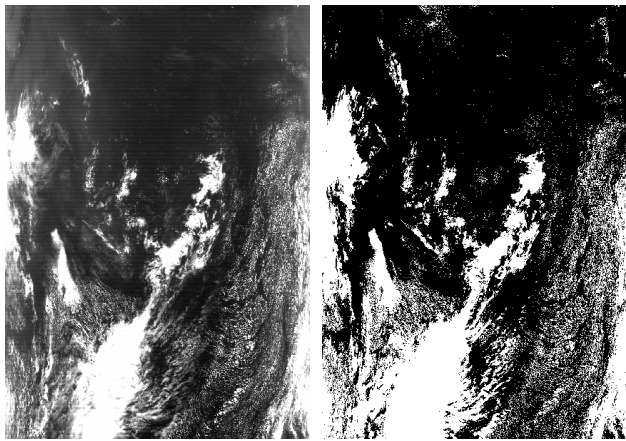


Figure 3. MODIS band 8 ($0.41\mu\text{m}$) and cloud mask.

IV. SUMMARY

A simplified cloud mask algorithm for onboard processing and data compression has been developed based on broadcast version of MODIS operational algorithm. The algorithm mainly uses visible/IR threshold method to separate cloudy scene from clear scene. The threshold varies with underlying surface scene types. In developing the algorithm, detailed analysis has been conducted to determine the performance of each test in the original algorithm by comparing the individual test results with MODIS operational cloud mask results. The algorithm bears 90% of accuracy compared with the original algorithm.

Clear sky data compression has been demonstrated with MODIS L1A and L1B data. Overall, the data volume reduction can be improved by over 40% with cloud mask extraction. The current cloud detection algorithm uses calibrated radiance/reflectance data currently requiring calibrations performed on ground. However, with advances in onboard processing and the availability of global positioning systems, both radiometric calibration and geo-location calibration can be performed on the spacecraft in the future. The eco-system data base can conceivably be supported also thus allowing cloud detection to be performed onboard.

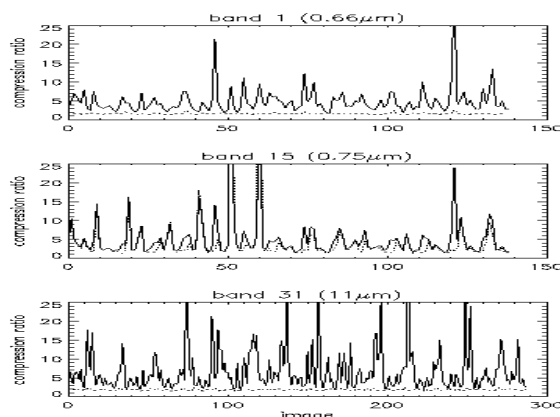


Figure 4. Compression ratios for selected MODIS bands from L1B data using cloud mask (solid lines) and without using cloud mask (dotted lines). Cloudy pixels are filled with the latest clear sky pixel value.

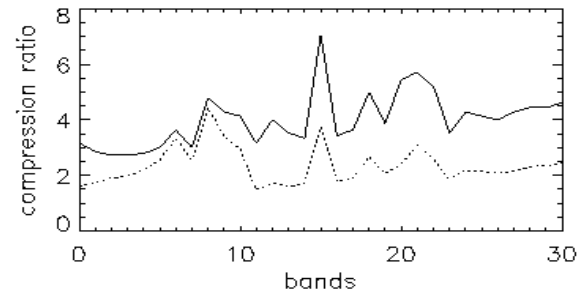


Figure 5. Compression ratios of L1A data using CCSDS lossless compression algorithm with cloud mask (solid line) and without cloud mask (dotted line). Results are based on 37 granules. Cloudy pixels are filled with the latest clear sky pixel value.

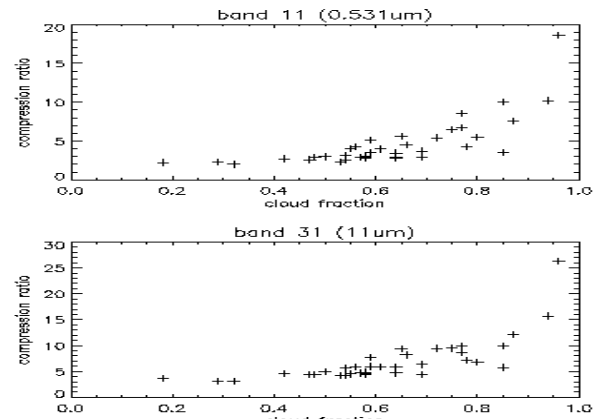


Figure 6. Compression ratio of L1A data versus cloud fraction. Cloudy pixels are filled with the latest clear sky pixel value.

- [1] Rossow, W. B. and L. C. Garder, 1993: Cloud detection using satellite measurements of infrared and the visible radiances for ISCCP. *J. Climate*, 6, 2341-2369.
- [2] Arking, A., 1985: Retrieval of cloud cover parameters from multispectral satellite images. *J. Climate Appl. Meteor.* 24, 322-333.
- [3] Kriebel, K. T., and R. W. Saunders, 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. *Int. J. Remote Sens.*, 9, 123-150.
- [4] Kriebel, K. T., R. W. Saunders, and G. Gesell, 1989: Optical properties of clouds derived from fully cloudy pixels. *Contr. Phys. Atmos.*, 62, 165-171.
- [5] Gesell, G., 1989: An algorithm for snow and ice detecting using AVHRR data: An extension to the APOLLO software package. *Int. J. Remote Sens.*, 10, 897-905.
- [6] Stowe, L. L., S. K. Vemury, and A. V. Rao, 1994: AVHRR clear sky radiation data sets at NOAA/NESDIS. *Adv. Space Res.*, 14, 113-116.
- [7] Ackerman S., K. Strabala, P. Menzel, R. Frey, C. Moeller, L. Gumley, B. Baum, C. Schaaf, and G. Riggs, 1997: Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35).
- [8] Hall, D. K., G. A. Riggs and V. V. Salomonson, 1995: Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data, *Remote Sensing of Environment*, 54, pp 127-140.
- [9] "Lossless data compression, CCSDS recommended for space data systems standards", Washington, DC, Blue Book, 120.0-G-1, 1997.
- [10] "Lossless data compression, CCSDS report concerning space data systems standards", Washington, DC, Green Book, 120.0-G-1, 1997.
- [11] Toller, G., A. Issacman, W. Barnes, MODIS Level 1B user's guide, MCST Document # Pub-01-0202-Rev B, MCST Internal Memorandum#M1038.
- [12] MODIS Tech. Rep. Series, NASA Tech. Mem. 104594, Vol. 2, GSFC, May 1994.