Postprint: An Automated Processing Method for All-Sky Camera Cloud Images

Authors: Wei Shiya[1]; Qiu Bo[1]; Cao Zihuang[2]; Fan Xiaodong[1]; Liu Yuanyuan[1]; Mu Yonghuan[1]

Date: 2018-06-22T00:00:00+00:00

Abstract

Utilizing all-sky cameras to capture cloud images is a widely adopted method in the astronomical community for monitoring sky cloud cover, and the resulting cloud cover estimation exerts significant influence on telescope observations. Currently, cloud cover estimation remains entirely manual, which is time-consuming, labor-intensive, and suffers from insufficient accuracy; the discrimination process is also completely dependent on personal experience. To address this issue, this paper proposes an automated calculation method for cloud cover specifically designed for all-sky camera images. The method first removes moon-affected regions from cloud images using temporal segmentation for cloudy images and differential methods for partly cloudy images. Subsequently, it performs binarization processing on the moon-removed cloudy images to segment clouds from the background, and employs a clustering algorithm based on grayscale values to quantitatively classify cloud thickness in partly cloudy images. Finally, it calculates the total cloud cover and automatically classifies the images according to the Thirty Meter Telescope (TMT) methodology for interpreting all-sky camera cloud images. Experimental results demonstrate that this method substantially improves the efficiency of cloud image interpretation, effectively reducing manual workload while achieving a relatively high recognition accuracy with an average of 76.67%.

Full Text

An Automatic Processing Method for All-sky Camera Cloud Images

Shiya Wei[1], Bo Qiu[1]*, Zihuang Cao[2], Xiaodong Fan[1], Yuanyuan Liu[1], Yonghuan Mu[1]
[1]Hebei University of Technology, Tianjin 300401

[2]National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012

Abstract

All-sky cameras are widely used in astronomy to monitor sky cloud conditions, and the subsequent estimation of cloud amount significantly impacts telescope observation planning. Currently, cloud amount estimation relies entirely on manual processing, which is time-consuming, labor-intensive, and lacks accuracy, with discrimination processes wholly dependent on personal experience. This paper proposes an automatic calculation method for cloud amount in all-sky camera images. The method first removes moon-affected regions from cloudy and partly cloudy images using time segmentation and difference methods, respectively. Then, it segments clouds from the background through binarization of cloudy images after moon removal, and quantifies cloud thickness in partly cloudy images using a gray-value-based clustering algorithm. Finally, total cloud amount is calculated and images are automatically classified according to the Thirty Meters Telescope (TMT) all-sky camera interpretation method. Experimental results demonstrate that this method substantially improves cloud image interpretation efficiency while effectively reducing manual workload, achieving an average recognition accuracy of 76.67%.

Keywords: All-sky cloud images; cloud amount; TMT

Introduction

Cloud observation and analysis constitute essential tools for atmospheric research and represent indispensable steps in astronomical observation. For instance, in astronomical site selection, only locations with predominantly clear skies throughout the year are considered suitable. Astronomical observations typically employ all-sky cameras to capture cloud images, from which the proportion of sky covered by clouds (cloud amount) is estimated. Following the TMT all-sky camera interpretation method, circles are drawn at zenith distances of 44.7° and 65°, marked in blue and green in [Figure 1: see original paper], dividing sky conditions into four categories: Clear (no clouds within the outer circle), Outer (no clouds in the inner circle but clouds between circles), Inner (clouds within the inner circle), and Covered (thick clouds covering over 50% of the area within the outer circle).

Historical literature has documented various automatic cloud image processing methods, though none follow TMT principles for automatic calculation. These approaches address two main aspects: first, cloud type classification, where features such as color, texture, and position identify cirrus, stratus, cumulus, etc., enabling rapid weather assessment and forecasting; second, cloud amount observation, primarily through remote sensing or infrared imaging. Zhou et al. achieved continuous daytime cloud monitoring by reselecting red-blue ratio thresholds for total sky imagers. Chen et al. calculated total cloud amount using whole-sky infrared images and statistical clear-sky thresholds for cloud pixel identification. Hu et al. obtained all-sky infrared radiation brightness temperature images through ground-based infrared cloud instruments, deriving cloud distribution and amount information via threshold segmentation, effectively reducing ground environment and solar illumination effects while enabling all-day operation. Cazorla et al. employed an optimized neural network classification with genetic algorithms to categorize images into clear sky and two cloud classes (thin and opaque), achieving only 61% accuracy for thin cloud detection. Li et al. proposed a thin cloud detection algorithm based on the MAP-MRF framework incorporating comprehensive features and spatial context, improving detection but remaining sensitive to illumination variations. Consequently, this research focuses on cloud detection under moonlight influence and classification of different cloud thicknesses.

This study addresses automatic cloud amount calculation and image classification according to TMT methodology. The processing workflow, shown in [Figure 2: see original paper], first removes moon-affected regions using time segmentation for cloudy images and difference methods for partly cloudy images. Subsequently, cloudy images undergo binarization to segment clouds from background, while partly cloudy images employ gray-value-based clustering to quantify cloud thickness. Cloud amount is then calculated separately for both categories, followed by automatic classification according to TMT interpretation standards.

1 Cloud Image Preprocessing

Given the variable nature of cloud image data, preprocessing is necessary to enhance algorithm robustness and calculation accuracy. Preprocessing comprises image enhancement and pre-classification.

1.1 Cloud Image Enhancement

Image enhancement is crucial for improving processing accuracy. This study employs morphological top-hat and bottom-hat transformations. The top-hat transformation acts as a high-pass filter, highlighting cloud gray-level peaks and enhancing cloud boundary information. The bottom-hat transformation detects valleys in cloud images, emphasizing boundaries of small or closely spaced clouds to prevent omission. Their combination further separates foreground and background gray levels, highlighting both detail and edge information while improving image contrast.

The morphological operations are defined as:

$$\mathrm{That}(f) = f - (f \circ b)$$

$$\mathrm{Bhat}(f) = (f \bullet b) - f$$

where f represents the original image, $\circ$ denotes opening operation, $\bullet$ denotes closing operation, and b is the morphological operator—a disk with a radius of 5 pixels in this algorithm. [Figure 3: see original paper] shows the original image (left) and enhanced image (right), demonstrating significantly improved cloud boundaries and enhanced small thin clouds.

1.2 Cloud Image Pre-classification

The dataset consists of 180 all-sky images captured continuously over two days using a Canon 550D camera with an EF 8mm f/2.8 fisheye lens, with 60-second exposure times and 5-minute intervals. Since moon influence varies across images—clearly distinguishable in partly cloudy images but partially or fully obscured in cloudy images—pre-classification improves processing precision. Histogram analysis reveals distinct characteristics: cloudy images exhibit higher pixel counts in the 100-150 gray-level range and lower counts in 0-50 compared to partly cloudy images, as shown in [Figure 4: see original paper]. This feature enables automatic classification of all images into cloudy and partly cloudy categories.

2 Removal of Moon Influence Regions

While preprocessing enhances cloud boundaries, moon-affected regions exhibit gray values similar to clouds, causing misclassification. This study addresses this through time segmentation for cloudy images and difference methods for partly cloudy images.

2.1 Time Segmentation Method

Time segmentation identifies moving objects and their trajectories based on minimal gray-value variation across consecutive frames. In cloud images, the moon moves significantly slower than clouds. Leveraging temporal continuity, this method compares each point' s gray value in a target image with corresponding points in 30 neighboring images.

The process involves: (1) identifying potential moon region points where the gray value difference between the target point and at least 30 corresponding neighboring points remains below a threshold, with the target point showing relatively high gray values; (2) verifying moon boundary points by checking for local gray-value maxima within the target point' s neighborhood, setting confirmed moon points to zero; (3) processing all points iteratively until complete moon identification.

While effective for moon removal in cloudy images, bright regions around the moon persist due to moonlight influence. After binarization, these high-intensity areas are easily misclassified as clouds in partly cloudy images. However, in cloudy images where surrounding clouds offset moonlight effects, cloud-background segmentation performs better, making this method more suitable for cloudy conditions.

2.2 Difference Method

Image subtraction, or the difference method, detects differences between same-scene images. Applied to cloud images, it exploits the moon' s positional invariance over short intervals to remove both moon and its influence region.

Observations reveal negligible moon position changes within 40 minutes. Therefore, 10 background images captured between midnight and 6:30 AM at approximately 40-minute intervals—containing moon but no clouds—were selected. The 180 target images, captured continuously over two days during the same time period, were preprocessed and subtracted from temporally matched background images, effectively removing moon influence and segmenting clouds from sky, as demonstrated in [Figure 5: see original paper]. This approach performs well when the moon is clearly visible and unobscured, making it suitable for partly cloudy images.

2.3 Comparison of Processing Effects

Applying difference methods to partly cloudy images and time segmentation to cloudy images yields distinct results. [Figure 6: see original paper] and [Figure 7: see original paper] show that difference methods remove moon influence while simultaneously segmenting clouds from background, whereas time segmentation effectively removes unobscured moon regions in cloudy images.

3 Cloud Amount Calculation

Following preprocessing and moon influence removal, clouds are effectively segmented from background, enabling separate calculation methods for each image category.

3.1 Cloud Amount Calculation for Partly Cloudy Images

In partly cloudy images, significant thickness variations exist across cloud regions. This study employs K-Means clustering based on gray values to analyze cloud thickness quantitatively. K-Means randomly selects k initial cluster centers, iteratively recalculates distances between samples and centers until cluster center displacement falls below a threshold, and classifies pixels based on gray-level variance features. Compared to alternative algorithms, K-Means offers faster convergence and superior clustering performance.

This research categorizes clouds into four classes—sky, thin cloud, middle cloud, and thick cloud—based on cluster center gray values, adjusting iteration counts to optimize classification. [Figure 8: see original paper] presents clustering results: the original image (left) and result (right) show cluster centers at gray values 36, 71, 0, and 104, corresponding to thin cloud, middle cloud, sky, and thick cloud, respectively. In the sky cluster (center=0), white represents background sky and black indicates cloud regions; in other clusters, white denotes clouds and black represents background.

International satellite cloud classification methods such as ISCCP, CLAVR-1, CLAVR-X, and MODIS calculate total cloud amount by classifying detected cloud pixels and computing probability-weighted sums. CLAVR-X categorizes pixels into thick cloud, mixed cloud, mixed clear, and clear sky. This study adapts CLAVR-X parameters, calculating total cloud amount for partly cloudy images using a weighted sum of four categories: thick cloud, middle cloud, thin cloud, and sky (with sky weight=0):

$$f(c) = \frac{0.9N_{thick\_cloud} + 0.6N_{middle\_cloud} + 0.3N_{thin\_cloud}}{N_{total}}$$

where N_{thick_cloud}, N_{middle_cloud}, and N_{thin_cloud} represent pixel counts for each cloud type, N_{total} is the total pixel count in the region, and $f(c)$ is the total cloud amount. Weight coefficients for middle and thin clouds are determined empirically based on their relative contributions.

3.2 Cloud Amount Calculation for Cloudy Images

For cloudy images where thick clouds dominate most regions with significant gray-value differences from background and minimal middle/thin cloud presence, this study classifies clouds into thick cloud and sky categories only. The adaptive Otsu thresholding algorithm performs binarization on moon-corrected cloudy images to segment clouds from background.

Otsu employs clustering principles to divide gray levels into two classes maximizing inter-class variance while minimizing intra-class variance, calculating an optimal segmentation threshold through variance analysis. This simple, robust algorithm operates independently of image brightness and contrast. Total cloud amount is calculated by counting all pixels with gray value=1:

$$f(c) = \frac{N_{cloudy}}{N_{total}}$$

where N_{cloudy} represents pixels with gray value=1, N_{total} is the total pixel count, and $f(c)$ is the total cloud amount.

3.3 Results Discussion

For method validation, 180 images were first manually classified to establish ground truth, then automatically classified using the proposed method. Comparison against manual classification yields recognition accuracies for each category, summarized in .

The results reveal discrepancies between manual and automatic classification for Inner, Clear, and Outer categories. Star points and acrylic dome reflections in all-sky images cause frequent misclassification of Outer and Clear as Inner. Specifically, 2 Outer images and 25 Clear images were misclassified as Inner, as illustrated in [Figure 9: see original paper] where white points inside the inner circle are stars and white regions represent dome reflections.

Overall, the proposed method achieves high recognition rates for Inner (93.9%) and Covered (88.46%) categories, though lower precision for Inner, Outer, and Clear due to noise sensitivity. The overall automatic recognition rate reaches 76.67%, meeting basic automatic processing requirements. In terms of efficiency, manual interpretation of 180 images requires approximately 30 minutes, while automatic processing completes in 2.1 minutes—a substantial improvement that becomes more pronounced with larger datasets.

This method offers rapid interpretation and satisfactory average accuracy. Since manual cloud amount observation lacks quantitative classification standards for Inner, Outer, and Clear categories, some deviation from manual results is expected. However, the 76.67% average accuracy across TMT categories satisfies automatic processing objectives.

The innovation lies in automatic removal of moon influence regions and quantitative cloud amount calculation/classification, effectively alleviating manual interpretation inefficiency. Current limitations include limited dataset size (180 images over two days) and fixed parameters. With expanded data, adaptive methods could optimize parameters—for instance, using maximum variance thresholding for moon detection threshold T, and determining cloud amount formula weights through statistical analysis of gray-level percentages across different cloud thicknesses. Future research should increase sample size, develop noise-specific models for effective removal, and specifically address misclassification of Clear as Inner under strong moonlight to improve overall accuracy.

References

[1] Peng Huanwen, Xin Yuxin, He Shousheng, et al. An Introduction to the All-Sky Camera at the YNAO Lijiang Astronomical Station [J]. Astronomical Research & Technology, 2015, 12(1):89-95.

[2] Heinle A, Macke A, Srivastav A, et al. Automatic cloud classification of whole sky images [J]. Atmospheric Measurement Techniques Discussions, 2010, 3(3):557-567.

[3] Zhou Wenjun, Niu Shengjie, Xu Xiaofeng, et al. Improvements of computational methods for cloud cover based on the total sky imager [J]. Transactions of Atmospheric Sciences, 2014, 37(3):289-296.

[4] Chen Lei, Han Yan, Qin Fangqiang, et al. Total Cloud Amount Calculation Based on Clear Sky Threshold Using Whole Sky Infrared Images [J]. Desert and Oasis Meteorology, 2015, 9(03):50-56.

[5] Hu Shuzhen, Ma Shuqing, Tao Fa, et al. An Experiment Study of All-sky Cloud Amount Observation Based on Infrared Real-time Threshold [J]. Quarterly Journal of Applied Meteorology, 2013, 24(2):179-188.

[6] Cazorla A, Olmo F J, Aladosarboledas L, et al. Development of a sky imager for cloud cover assessment [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2008, 25(1):29-39.

[7] Li Q, Lu W, Yang J, et al. Thin Cloud Detection of All-Sky Images Using Markov Random Fields [J]. IEEE Geoscience & Remote Sensing Letters, 2012, 9(3):417-421.

[8] ZHU Shihu. Function extension and application of morphological top-hat transformation and bottom-hat transformation [J]. Computer Engineering and Applications, 2011, 47(34):190-192.

[9] Wang Enwang, Wang Enda. Application of an Improved Frame Difference Method in Space Moving Target Detection [J]. Astronomical Research & Technology, 2015, 17(5):87-89.

[10] Arghavan Keivani, Jules-Raymond Tapamo, Farzad Ghayoor, et al. Motion-based moving object detection and tracking using automatic K-means clustering [J]. IEEE Africon 2017 Proceedings, 2017, 32-37.

[11] Maddy E S, King T S, Sun H, et al. Using MetOp-A AVHRR Clear-Sky Measurements to Cloud-Clear MetOp-A IASI Column Radiances [J]. Journal of Atmospheric and Oceanic Technology, 2011, 28(28):1104-1116.

Note: Figure translations are in progress. See original paper for figures.