Two approaches for automated cloud detection from telescope data are investigated. The first approach exploits variability of source flux as a detection mechanism. The second approach exploits the spatio-temporal variability of clouds to discriminate them from stellar background. Three algorithms in the second category, based on temporal, statistical and spatial variability as discriminating features, are described and preliminary results are presented. For the spatial approach, performance evaluation with hand-truthed telescope data is presented.

1. INTRODUCTION

Clouds can degrade the capabilities of ground based space surveillance optical sensors by occluding targets of interest. Various approaches for cloud detection are summarized in Table 1. These options range from visual observation, to dedicated IR sensors for cloud detection. Both are expensive approaches, as ancillary sensor support and/or personnel overhead are required.

<table>
<thead>
<tr>
<th>Class of Options</th>
<th>Method/System</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Observation</td>
<td>Visual inspection of search sectors multiple times on hour / re-task as appropriate</td>
<td>Timely</td>
<td>Personnel overhead</td>
</tr>
<tr>
<td>Weather Data</td>
<td>Local forecast / weather station data</td>
<td></td>
<td>Not necessarily concurrent with collection tasking</td>
</tr>
<tr>
<td>Dedicated Sensor</td>
<td>Visible / NIR</td>
<td>Detects clouds by airglow extinction</td>
<td>Never demonstrated Airglow variable and has spatial structure</td>
</tr>
<tr>
<td></td>
<td>Non-Imaging M/NIR</td>
<td>No moving parts</td>
<td>Required temperature differential for detection (High, cold clouds)</td>
</tr>
<tr>
<td></td>
<td>Direct detection of clouds in LWIR</td>
<td>Simple processing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simple electronics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No enclosure</td>
<td></td>
</tr>
<tr>
<td>Data Processing Options</td>
<td>Image M/NIR / LWIR</td>
<td>No moon occultor</td>
<td>Required temperature differential</td>
</tr>
<tr>
<td></td>
<td>Direct detection of clouds in LWIR</td>
<td></td>
<td>Rotational complex system calibration</td>
</tr>
<tr>
<td></td>
<td>Maintains exclusion zone map</td>
<td>Inexpensive, no dedicated sensor required</td>
<td>Must miss track to detect exclusion</td>
</tr>
<tr>
<td></td>
<td>Serendipitously as system misses tracks</td>
<td></td>
<td>Regions “go clear” in time, even if still cloudy</td>
</tr>
<tr>
<td></td>
<td>Algorithms for cloud detection</td>
<td>Inexpensive, no dedicated sensor required</td>
<td>Detection based on assumed cloud spatial/ temporal/ statistical behavior</td>
</tr>
</tbody>
</table>
With the increasing field of view (FOV) and search rates of future optical sensors, cloud impact increases. However, at the same time, cloud structure and temporal variability becomes discernable in the surveillance telescope data, and gross estimates of cloud cover can be replaced with a finer assessment and extraction of usable data (See Fig. 1). The impact is two-fold. First, photometric and astrometric inaccuracies can be identified and linked to specific data sets. Second, data can be further segmented, identifying portions that are cloud covered prior to processing, to improve detection and tracking performance.

We investigate two data processing options that exploit the optical imagery data from space surveillance telescopes for data image quality assessment and generation of cloud maps. The first approach leverages stellar extinction to generate cloud masks. Flux from star detections are compared to estimated average fluxes on cloud free nights. Areas with significant flux change are identified as locations of potential clouds. The second approach takes advantage of structural and temporal variability to discriminate clouds from the stellar background. Three algorithms for automated cloud map generation were developed. The first algorithm leverages multi-frame data to generate a statistical background model of the stellar field. Pixels that vary significantly from the expected background model are identified as clouds. For the temporal approach, an optical flow algorithm was employed to assess per pixel relative frame to frame motion. Non-stationary pixels are identified as clouds. The last algorithm uses texture discrimination to discriminate cloud from stellar background. Gabor wavelet decomposition is used to derive a multi-dimensional feature vector for each pixel in a frame. Matched filtering using a stellar background centroid feature vector derived from a training set is used to identify cloud regions.

Results will be presented from data from the Lincoln Near-Earth Asteroid Research (LINEAR) Ground-based Electro-optical Deep Space Surveillance (GEODSS) telescope and from a wide field of view camera deployed alongside the Experimental Test Site (ETS)-B in Socorro, New Mexico.

2. DATA

LINEAR is an MIT Lincoln Laboratory program to detect and catalogue near-Earth objects (NEOs) (see Fig. 2). The project uses two of the Ground-based Electro-Optical Deep Space Surveillance (GEODSS) telescopes, each are equipped with Lincoln developed CCD electro-optical detectors. Each sensor has a two square degree field of view, and a search rate of 1200 degree square per night. Four to seven sectors or 600 fields are surveyed each night, at a revisit rate 20 minutes and an integration time ranging between 4 – 5 seconds. There is limited truth data available at site. A binary cloud cover assessment is made by an operator during collection and is recorded in a database. For the study, this assessment as well as cloud masks for 260 frames generated by visual inspection of the image data, are used as truth masks for algorithm evaluation. A sample frame and estimated cloud truth mask is shown in Fig. 3.

The temporal technique was tested using the camera attached to ETS-B. This sensor is a wide FOV system, 5.5 x 5.5 sq degrees, with a sampling rate of 0.28 seconds between frames. This was a sufficient sample rate to detect apparent cloud motion across the stellar background.
3. METHODS

Each of the algorithmic approaches investigated for development of cloud masks is dependent on the assumed characteristics of the cloud structure that allows discrimination from the background stellar field. The algorithm flow diagrams as well as sample results for each method are shown in Fig 2(a)-(d). For all data sets, the image data is first preprocessed by dynamic range rescaling for contrast enhancement by using the following nonlinear remapping function [1].

\[ I_0 = \log(1 + I_i) \]

, where \( I_i \) is the input image, \( I_0 \) the output image.
The first approach leverages the change in measured stellar flux due to the presence of clouds. The Lincoln Laboratory developed Optical Processing Architecture (OPAL) was used to estimate star positions from the mean images of several image frames collected from a single field. For select star positions, the estimated average flux over these image frames was calculated. For a new frame of data, star positions for which there was 200 counts/second or more flux change was labeled as potential location of cloud cover. The resultant image was divided into 40 x 40 pixel blocks. If a star position was associated with significant flux change, the block is identified as a cloud block. The algorithm plus result is shown in Fig. 4(a).

The temporal approach (Fig. 4(b)) discriminates clouds based on the movement of structure across the field of view. Because the revisit time of 20 minutes for LINEAR was not sufficient for tracking cloud motion, the camera attached to ETS-B was used. Ten minutes of data at time sampling of 0.28 minutes was used. Each pair of consecutive frames was used as inputs to the Lucas-Kanade [2] optical flow algorithm. The resulting speed map was threshold for relative speeds > 1 pixel/ (sample time). The result is a series of cloud map images showing the time evolution of the cloud across the focal plane. Fig. 4b shows one frame from a movie of the evolving cloud map.

The statistical approach (Fig. 4(c)) uses multiple frames of registered data to generate a stellar background model [3]. These data sets may or may not include frames with cloud present. In this implementation, the per pixel co-registered statistics are modeled as Gaussian distribution. The outputs are two images: a mean and standard deviation. Because of camera jitter, image pixels were aggregated into 2x2 blocks prior to statistics calculation. A test image is normalized by performing a point to point subtraction of the mean image and division by the standard deviation image. Large pixel values are then associated with cloud maps.

The last method exploits the textural dissimilarity between cloud and stellar background (Fig. 4(d)) [4]. A single input frame is input to a filter bank of 2D Gabor wavelets.

$$G(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cos\left(\frac{2\pi x}{\lambda}\right)$$

where $\sigma$ is the variance of the Gaussian function, $\gamma$ is the ellipticity, $\lambda$ is the scale, $\theta$ is the orientation angle and the transformed coordinates at the rotation angle are given by:

$$x' = x \cos \theta + y \sin \theta$$
$$y' = -x \sin \theta + y \cos \theta$$

In this implementation, 4 angles between 0 and 135° and 7 frequencies at 1 octave separations were chosen for the Gabor wavelet filter bank. After filtering the image with the 28 filters in the wavelet filter bank, the result is passed through a sigmoidal nonlinearity function to generate 28 multi-scale blob images. The registered set of blob images is then used to generate per pixel feature vectors. These are 31 component feature vectors, 28 comprised of the output of texture output, the associated $x$ and $y$ pixel coordinate and the original image pixel amplitude. Each pixel feature vector is then normalized to have mean of 0 and standard deviation of 1. A stellar feature vector is derived separately from known cloud free data. For computational speed, a fraction (6%) of the pixels are first selected for matched filtering A matched filter is then used to assess angular separation of each feature vector to the stellar feature vector and the image is segmented based on an angular threshold. The remaining pixels are classified using the Mahalanobis distance:

$$D(x, y) = \sqrt{(x - y)\Sigma^{-1}(x - y)}$$

where $x$ and $y$ are two feature vectors, and $\Sigma$ the covariance matrix.

For all four techniques, the resultant cloud detection map is dependent on the selected threshold, and thus performance can be evaluated by generating an associated receiver operating characteristic (ROC) curve. In the next section, the ROC curve for the texture based cloud detection method will be presented.
Fig. 4. Anomaly detection approaches to cloud detection. (a). Changing stellar flux. (b) Temporal variability via optical flow. (c). Temporal variability via background modeling. (d). Spatial texture scale.
4. RESULTS

The original code for the texture based algorithm was written in MATLAB and run on the LLGrid cluster at Lincoln Laboratory. LLGrid is comprised of 140 dual-processor compute nodes (280 processors). Each node consists of 2 Intel Xeon processors (3.2 GHz) and 6GB RAM. Of this, 32 processors were available for processing data. The elapsed time for running 57 files on 32 processors is approximately 40 minutes. For 260 files, the run time was approximately 2.5 hours.

Fig 5. shows the percent detection result for two nights of LINEAR data and compares it to the binary operator assessment of cloud cover recovered from site. For nights that were labeled as clear, the detection algorithm indicates cloud cover of less than 5%, consistent with operator reporting. For nights labeled as cloudy, the automation algorithm shows mostly 100% cloud cover. There are however frames that show less than 50% cloud detection on some instances. This indicates that there are available frames that are cloud free even in operator labeled cloudy conditions that may contain useful data. These frames can be analyzed in more detail by comparing the resulting detection map at a fixed angular threshold to the user assessed truth cloud mask.

Fig. 6 (a) shows sample cloud maps output from the algorithm. For each of the 250 frames, the percent pixel overlap with truth data and the percent false alarm for a fixed threshold was calculated. The result for all angular thresholds is shown in the ROC curve in Fig. 6 (b). The knee of the curve shows that for a 70% detection rate, there is a false positive rate of 20%.

There are two components to mismatch between the estimated cloud detection map and the truth mask generated by visual inspection. First, is the subjective nature of cloud truthing. Typically, the estimated truth mask includes areas of cloud peak-through whereas the texture based method works optimally on opaque clouds. The automated cloud masks therefore tend to be slightly smaller in extent compared to the operator defined truth masks as shown in Fig 6 (a). Most of the false alarms came surprisingly from the cloud free nights, where the cloud detection algorithm picked regions near the image edges as components of clouds as well as shown in Fig 6 (b). These false alarms due to edge effects that show up in the sigmoidal nonlinearity function outputs. This can be mitigated by first extended the image past its boundary by mirroring local pixels prior to processing.
The second potential contributor is the feature vector that was used as the typical stellar background. Typically, the feature vector should be dependent on a selected field. A larger sampling of the historical LINEAR data for generation of the feature vector should moderately improve performance.

5. CONCLUSIONS

Four separate approaches exploiting resident telescope data for cloud detection have been highlighted and preliminary performance assessment for the textural based technique was presented. These methods take advantage of the spatio-temporal characteristics of cloud versus the stellar background to discriminate between the two and automatically generate a cloud mask for downstream quality assessment and data segmentation. Exploiting resident optical data is inexpensive compared to maintaining a separate auxiliary sensor designed specifically for cloud detection, and with sufficient computing power, may be done in a timely manner prior to down-stream processing. However, their performance is dependent on a priori knowledge of cloud spatial variability and dynamics.

Initial performance assessment shows promising results for the automatic detection based on textural discrimination and possible methods to improve performance was discussed. The remaining methods will be assessed in a similar manner and a methodology to combine one or more schemes for optimized detection is being investigated.

6. ACKNOWLEDGEMENT

The authors would like to thank Richard Lambour, Scott Stuart and Jenifer Evans for their discussion of the LINEAR system, as well as Brian Shucker and Michael Rodehorst for their help with OPAL and the flux-based cloud discrimination algorithm. This work is sponsored by the Air Force under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.

7. REFERENCES


Fig. 6. Performance assessment for texture based cloud discrimination. (a). Estimated cloud map for angular separation of 10°. (b) ROC curve (in blue) for 250 truthed images. In red, the 50-50 performance curve for comparison. (To be augmented with more data).