SANDIA REPORT

Extraction of Cloud Statistics from Whole Sky Imaging Cameras

W. Philip Kegelmeyer, Jr.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

SF 2900Q(8-81)

This report has been reproduced from the best available copy.

Available to DOE and DOE contractors from:

Office of Scientific and Technical Information
P. O. Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from:

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.
Springfield, VA 22161

EXTRACTION OF CLOUD STATISTICS FROM WHOLE SKY IMAGING CAMERAS

W. Philip Kegelmeyer, Jr.

Scientific Computing Department

Sandia National Laboratories/California

Abstract

Computer codes have been developed to extract basic cloud statistics from whole sky imaging (WSI) cameras. This report documents, on an algorithmic level, the steps and processes underlying these codes. Appendices comment on code details and on how to adapt to future changes in either the source camera or the host computer.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

ACKNOWLEDGMENT

Contents

1 Introduction

Computer codes have been developed to extract basic cloud statistics from whole sky imaging (WSI) cameras. The intent of this report is to document the steps and processes underlying these codes at an algorithmic level. (The source code and its comments is intended to serve as documentation at the more detailed implementation level.)

This report begins, in the next section, with a summary of the entire processing chain. The two subsequent sections treat, in more detail, the required pre-processing and the extraction of statistics proper. Two appendices comment on code details and on how to adapt to future changes in either the source camera or the host computer.

2 Summary

A terse summary of the overall process follows, and is elaborated upon in the following sections:

Input Image Description. The input to the cloud statistics algorithm is expected to be a grey-scale "cloud decision" image, as generated by the Scripps Marine Physical Laboratory from a WSI model EO-5 camera [6] field image.

> Points to note are that, as a WSI image it portrays a fish-eye projection of the sky, and as a "cloud decision" image each of its pixels represents, not a radiance value, but some value indicating the cloudiness of the sky at that point.

Pre-Processing. Accordingly, some pre-processing is required before cloud statistics can be extracted. The input image is processed by the Pseudo-Cartesian Transformation (PCT) [3] to convert it to a standard orthographic projection. Then thresholds (from [4]) are applied to the grey-scale cloud decision values to generate a processed image in which every pixel has a label representing one of three values: BACKGROUND, CLEAR, or CLOUD. This process can result in spurious single pixel regions near curved parts of the background; a minimum/maximum size filter [5] is applied to fill these in without changing the contours of any larger objects.

> Finally (and only after the "global" statistics have been extracted; see the next paragraph and section 4), any cloud touching the background region is erased. Such a cloud is only partially visible due to being clipped by the edge of the field of view, and would thus generate inaccurate statistics; it is therefore erased before per-cloud statistics are extracted.

Statistics: Two types of statistics are gathered, "global" and "per-cloud". The global statistics are computed over the entire original image, and the "per-cloud" statistics are computed only over clouds that are entirely visible (and so survive the "discard-partial-clouds" step).

> **Global:**
> - Fractional cloud cover; the percentage of visible sky pixels that are cloud pixels.
> - Mean and standard deviation of cloud path run lengths (defined below), averaged over the total number of such runs.
> - Mean and standard deviation of clear sky path run lengths (defined below), averaged over the total number of such runs.

Per Cloud:

- The number of fully visible clouds.
- Mean and standard deviation of cloud area, averaged over the number of clouds.
- Mean and standard deviation of cloud perimeter, averaged over the number of clouds.
- Mean and standard deviation of a shape measure that expresses the "circularity" (or lack thereof) of a cloud, averaged over the number of clouds.

All statistics with dimensions are measured in image pixels. If the input image had a cloud field of relatively uniform height, and that height is known and provided to the algorithm, then the size of a pixel can be computed. In this case the dimensioned statistics will be reported in the same units as the provided height.

3 Pre-Processing

3.1 Pseudo-Cartesian Transformation

The WSI camera provides a fish-eye view of the sky. This is useful in that it allows the imaging of the entire sky dome, but it also has the effect of altering cloud shapes. In the WSI view, radial distance from the center of the image is proportional to zenith angle. Accordingly, pixels near the edge of an image subtend a smaller absolute distance than those directly overhead, with the result that a particular cloud will appear expanded or compressed according to whether or not the WSI is directly beneath it or off to one side. An example WSI cloud decision image is in Figure 1.

For the shape of a cloud to be independent of zenith angle, the WSI image must be "flattened" by applying the pseudo-cartesian transformation (PCT), [3]. The technical details of our implementation are discussed at length in [1, Chapter 4]; the basic idea is that the PCT rescales distance from the center of the image to a dependence that varies linearly with the tangent of the zenith angle. The main practical point to be aware of is that the PCT must choose some subarea of the original field of view to process. Given the increasing distortion of the WSI camera with increasing zenith, we here apply the PCT to only the field of view down to 65° off zenith. Thus in the output, only the sky dome and clouds present in the first 65° are visible and available for subsequent statistics extraction.

An example of the PCT, as applied to the original image in Figure 1, is in Figure 2.

3.2 Cloud Decision Thresholding

As discussed in detail in [4], the cloud decision images as generated by MPL from the raw WSI images have pixel values which encode the nature of the sky at that point, generally according to its "cloudiness". The mapping is as follows:

0:   Background; the pixels outside of the camera's field of view, or hidden by the sun occluder arm.
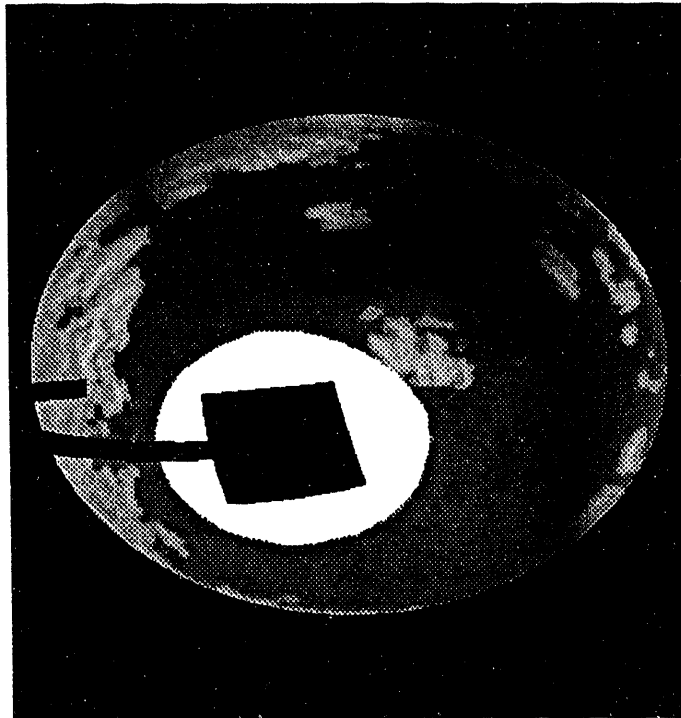
1–99:   Clear sky pixels.

7

Figure 1: An Input WSI Cloud Decision Image.

100–139: Thin cloud.

140–200: Opaque cloud.

201–255: Off scale, artifacts of the cloud decision process which cannot be classified. This is primarily the "halo" around the occluder, which cannot be classified as sky or cloud.

For the purposes of gathering statistics, we need to know, for each pixel, only whether it is non-sky background (the parts of the square image outside of the round field of view of the WSI), clear sky, or cloud. Therefore we apply a series of thresholds so that the resulting image has only three values (given symbolic names here, but they are integer values in the image):

BACKGROUND: All pixels with value 0 or 201 and higher in the PCT'ed image.

CLEAR: All pixels with values in [1, 99], inclusive, in the PCT'ed image.

CLOUD: All pixels with values in [100, 200], inclusive, in the PCT'ed image.

The thresholding process can reveal the existence of a handful of spurious single pixel regions near highly curved parts of the background and the occluder. These artifacts are not due to meteorological phenomena; they occur because of the digital nature of the image, and the fact that it is difficult to render a smooth curved edge in blocky pixels. To eliminate these single pixel regions without changing the shape or contours of any larger cloud region, a morphological size
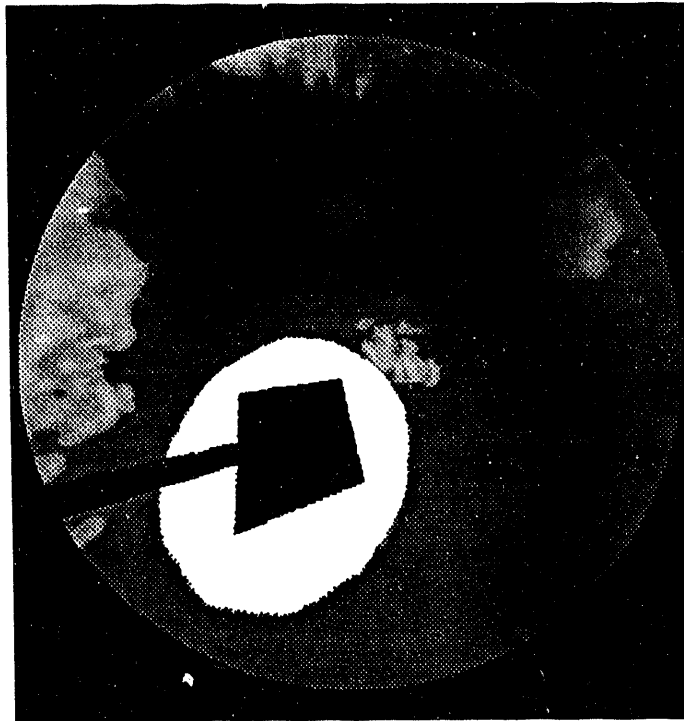
8

Figure 2: The Result of Pseudo-Cartesian Transformation.

filter is applied, [5]. Basically, every pixel is replaced by the minimum of its 3 by 3 neighbors in the first pass of the filter, and then by the maximum of its 3 by 3 neighbors in the second. This operation leaves the interior of regions and the contours of regions larger than a single pixel unaffected, but merges single pixel regions into the dominant surrounding region.

An example of the three-valued cloud decision image that would be extracted from Figure 2 is in Figure 3.

3.3   Removal of Partial Clouds

The global statistics (see Section 4) are extracted on the PCT'ed and decision thresholded image without any further processing. However, all of the other statistics to be extracted depend on cloud size or shape, and, as can be seen from any of the figures so far, there are clouds that exist at the edge of the field of view or at the edge of the occluder arm, and so which are not entirely visible to the camera. Thus their entire size and shape cannot be accurately accessed, and so statistics extracted from such partial clouds would be suspect.

Accordingly, these partial clouds are detected and erased (that is, given the value BACK-GROUND) before the per-cloud statistics are computed. Figure 4 demonstrates the effects of removing partial clouds from Figure 3.

9
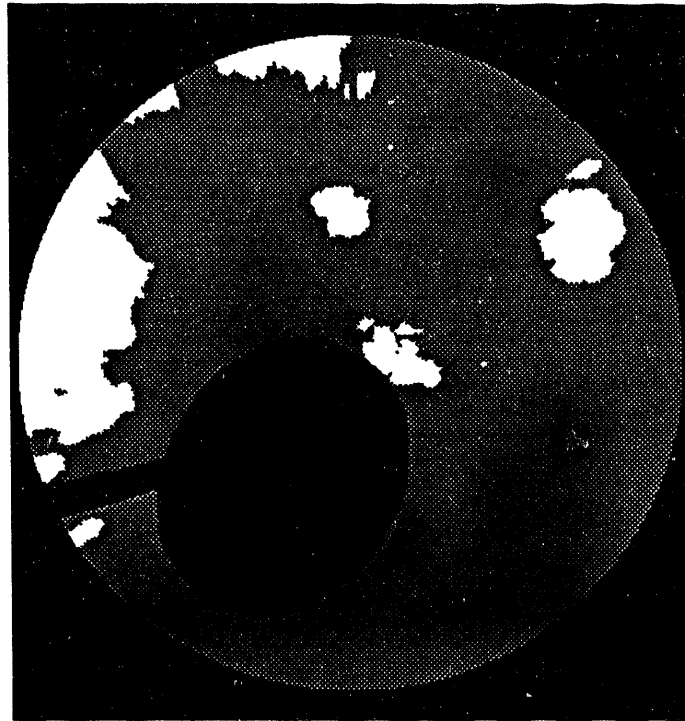
Figure 3: A Thresholded Cloud Decision Image.

4 Statistics Computation

Two sorts of statistics are extracted. The "global" statistics are computed from the PCT'ed thresholded cloud decision image *before* removal of the partial clouds, and the "per-cloud" statistics are computed after the removal. Notes on definition and implementation of each statistic follow.

4.1 Global Statistics

Cloud Cover Fraction: This is computed as the number of pixels of type CLOUD over the total of the number of pixels of type CLOUD or CLEAR.

Cloud Path Length, Mean and Standard Deviation: The image is scanned both horizontally and vertically. In each row or column, each "run" of contiguous CLOUD pixels is noted and its length recorded For example, if 0 is clear sky and X is cloud, then the following row of sky pixels

0 0 0 0 0 X X X X X X X X 0 0 0 0 0 0 X X X X 0 0 0 0 0 0 0 0 0 0 0

would be considered to contain two CLOUD runs, one of length eight and one of length 4. Once all horizontal and vertical CLOUD run lengths are recorded, their mean and standard deviation are computed.

Figure 4: A Thresholded Cloud Decision Image with Partial Clouds Removed.

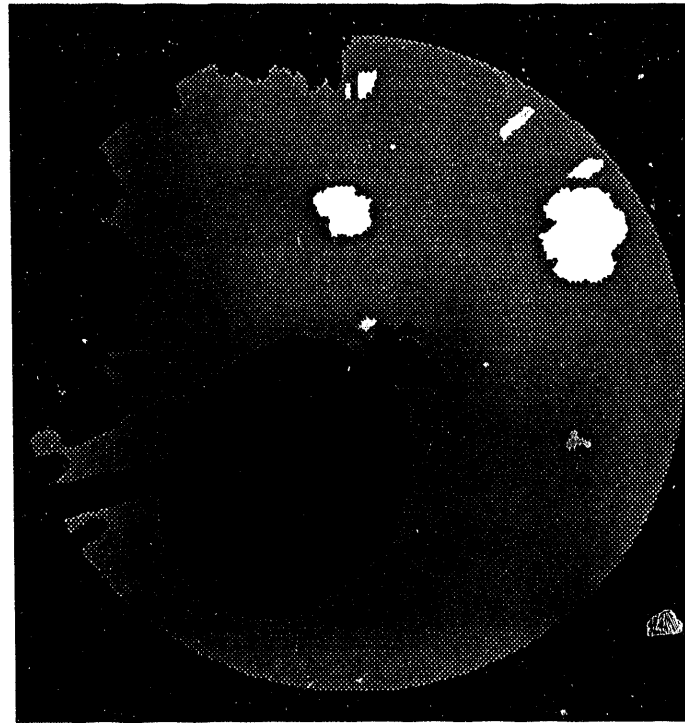**Clear Path Length, Mean and Standard Deviation:** Clear path length is computed similarly, except that runs of CLEAR pixels are noted. The example above contains three such runs, of lengths five, six and ten.

4.2   Per-Cloud Statistics

Number of clouds: Any contiguous region of four-connected (that is, connected vertically or horizontally) CLOUD pixels is considered to be a separate cloud. The number of such clouds is recorded.

Cloud Area, Mean and Standard Deviation: The area of each cloud is defined as the number of CLOUD pixels it contains. The mean and standard deviation of the areas is computed over the number of clouds.

Cloud Perimeter, Mean and Standard Deviation: The perimeter of each cloud is defined as the number of CLOUD pixels it contains which have a horizontal or vertical neighbor which is a SKY pixel. The mean and standard deviation of the perimeters is computed over the number of clouds.

Cloud Circularity, Mean and Standard Deviation: The "circularity" of each cloud is defined here as
$P^2/(4\pi A)$, where P is perimeter and A is area. This value is known as the isoparametric ratio (derived from the isoparametric inequality, [2]), and has the property that it is provably always

11

greater than or equal to 1, with equality only if the object is a perfect circle. Accordingly, it is commonly used as a rough indication of an object's shape. The mean and standard deviation of the isoparametric ratio is computed over the number of clouds.

4.3   Units of Measurement

All statistics with dimensions are originally measured in image pixels. If the input image contains a cloud field of relatively uniform height, and that height is known and provided to the algorithm, then the size of a pixel in absolute units, in the PCT'ed image, can be computed relatively easily. The resultant formula is

$$\text{width of pixel} = \left(\frac{\tan(65°)}{235} \right) \times \text{height of cloud field,} \tag{1}$$

given the PCT parameters used here.

Thus if the height is known and provided to the algorithm, the dimensioned statistics are multipled by the above factor when they are recorded, so that they will be reported in the same units as the provided height.

A Comments on the Source Code

The code files of primary interest are `cloudstats.h` (which contains the definition of the cloud statistics structure and some defined constants) and `cloudstats.c`, which contains the main functional interface, `cloud_characteristics()` and all of the statistics related support code. The code is thoroughly internally commented, and should be considered the first resource for further details on any implementation or definition issue.

The files `defs.h`, `main.c` and `wsi.c` support the computation of the PCT, and thus may be of some interest as well. The file `dummy_main.c` is simply an interface to `cloud_charact-eristics()`, and `lvxvtorip.c` contains a function to support conversion between RIPPLE and Khoros image formats; neither file is likely to have much intrinsic interest.

B Comments on Computer or Camera Changes

B.1 Operating System Dependencies

The current release of the cloud statistics software has three components; C source code from Sandia, a dependence on some of the Khoros include files and object libraries (not included but separately available without fee; send e-mail to `khoros@chama.eece.unm.edu` for further information), and a dependence on the RIPPLE include files and object libraries (included). The code has been developed under the SunOS 4.1.3 operating system, and will port without modification to any other recent, but pre-Solaris, Sun operating system.

If the code is to be ported to a non-Sun OS, or upgraded to Solaris, then the appropriate Khoros release will be needed. Further, the RIPPLE object libraries will have to be ported to the new system by Sandia Laboratories. This could be trivial or difficult, depending on whether Sandia already supports the architecture and OS in question internally. Inquiries to the author of this report are welcome.

B.2 Camera Dependencies

The algorithms described in this report were developed for images from the MPL model EO-5 WSI using the cloud decision algorithm described in [4]. If the camera model or the cloud decision algorithm changes, then changes in the code will be required as well. In the former case, consult [1, Appendix B.4] for more direction; the basic idea is that some defined constants will have to be changed in `defs.h` to permit the PCT to be properly computed, and it may be necessary to change the functions in `wsi.c` which implement the camera calibration equation.

In the latter case (changes in the cloud decision algorithm), the changes to the code will have to depend on how the new algorithm behaves. Keep in mind that the goal is a ternary image with every pixel labeled CLOUD, SKY, or BACKGROUND. This might be achievable by changing the threshold constants in `cloudstats.h`, or it might require some new code. Once that ternary image is created, however, the rest of the cloud statistics software should proceed as before.

B.3 Thin vs. Thick Clouds

The current algorithm groups thin and thick clouds into one class and measures statistics on them all at once. If statistics on just thick clouds or just thin clouds were desired, this could be accomplished with the current cloud decision images, again by simply changing the threshold constants in cloudstats.h under the direction of [4], and recompiling.

Note, though, that since this method essentially just erases the type of cloud not of interest, some of the thin clouds processed might be just the perimeter of a normal cloud, and so will look like a "halo". If it is undesirable to extract statistics on such halos, then extra code will have to be written to detect such artifacts and either erase them or fill them in.

C Bibliography

[1] ALLMEN, M., AND KEGELMEYER, JR., W. P. The computation of cloud base height from paired whole sky imaging cameras. SAND report, Sandia National Laboratories, MS 9214, P.O. Box 969, Livermore, CA, 94511, February 1994.

[2] DoCARMO, M. *Differential Geometry of Curves and Surfaces.* Prentice-Hall, 1976.

[3] KOEHLER, T., AND SHIELDS, J. Factors influencing the development of a short term CFARC prediction technique based on WSI imagery. Technical Note 223, Marine Physical Laboratory, SCRIPPS Institution of Oceanography, San Diego, CA, 92151-6400, November 1990.

[4] MARINE PHYSICAL LABORATORY. CMPDEC image and header format. Technical Memo #AV 91-031t, SCRIPPS Institution of Oceanography, UC San Diego, 1991. Contact Monette Karr, (619) 534-1767.

[5] PRATT, W. K., AND RISSER, T. B. Digital processing of medical imagery. *SPIE Applications of Digital Image Processing 359*, 4 (1982), 156–162.

[6] SHIELDS, J., KOEHLER, T., AND KARR, M. Automated cloud cover & visibility systems for real time applications. Technical Note 217, Marine Physical Laboratory, SCRIPPS Institution of Oceanography, San Diego, CA, 92151-6400, January 1990.

UNLIMITED RELEASE
INITIAL DISTRIBUTION

MS 0735   D.A. Arvizu, 6200

          Attn: C.P. Cameron, 6215
                     C.E. Tyner, 6216
                     D.E. Hasti, 6218

MS 0722   D. Emgi, 6904

MS 9001   J.C. Crawford, 8000

          Attn: D.L. Crawford, 1900
                     E.E. Ives, 5200
                     J.B. Wright, 5300
                     R.J. Detry, 8200
                     W.J.McLean, 8300
                     L.A. Hiles, 8400
                     P.N. Smith, 8500
                     L.A. West, 8600
                     R.C. Wayne, 8700
                     M.T. Dyer, 8800

MS 9004   M. E. John, 8100

MS 9103   G.A. Thomas, 8111

MS 9201   R.M. Wheeler, 8112

MS 9031   J.S. Swearengen, 8113

MS 9201   R.J. Gallagher, 8114

MS 9104   M.H. Rogers, 8115

MS 9202   R.L. Bierbaum, 8116

MS 9056   J. Vitko, Jr., 8102

MS 9056   L.R. Thorne, 8102

MS 9056   C.H. Sun, 8102

MS 9056   M. Lapp, 8102

MS 9214   L.M. Napolitano, 8117

MS 9214   M.C. Allmen, 8117

MS 9214   W.P. Kegelmeyer, Jr., 8117 (10)

MS 9014   K.A. Buch, 8351

MS 9021   Technical Communications for OSTI (10)

MS 9021   Technical Communications/Technical Library Processes, 7141

MS 0899   Technical Library Processes (4)

MS 9018   Central Technical Files (3)

DATE FILMED
5/12/94

END