Fractals, Vol. 2 No. 3 (1994) 399–404

QUANTITATIVE CLASSIFICATION OF CLOUD MICROPHYSICAL IMAGERY VIA FRACTAL-DIMENSION CALCULATIONS

L. IAN LUMB, TOM B. LOW and ARTHUR DI LEO
KelResearch Corporation
850-A Alness Street, Suite 9
Downsview (Ontario), M3J 2H5, Canada
e-mail: `kel@nexus.yorku.ca`

The use of fractal-dimension calculations, for quantitative classification of various objects, is well established in many areas of the physical and life sciences. Such fractal-dimension calculations are useful in that they furnish some measure of geometrical complexity that is not available through "traditional" approaches. In the present context, we focus attention on the application of such calculations to cloud microphysical data. In particular, we consider data from *in-situ*, two-dimensional particle imagery. Whereas previous efforts have typically characterized the imaged hydrometeoric fields by dimensional and statistical measures, distinguishing between the various hydrometeoric types should also be possible with fractal dimension based analyses. Although typical data sets may include thousands of individual images, each of these single images is typically quite small in size, and has intensity values that span only a few levels. Thus the need for computationally efficient algorithms, that can process these individual images, presents interesting challenges for fractal-dimension calculations. Here we report on our preliminary findings regarding the capacity and information dimension of various synthetic images.

1. OVERVIEW

On synoptic (2000-5000 km) to planetary length scales, clouds comprise an important component of the atmospheric system, and are thus involved in weather system to global cycling scale processes. Although individual clouds are mesoscale (5-20 km) features, efforts aimed at understanding cloud characteristics and dynamics in more detail must consider processes that take place on the "subcloud" scale, i.e. the very scale on which the formation of precipitation takes place. On this *microphysical* scale developing hydrometeors, which are themselves products of the condensation or sublimation of atmospheric water vapor, exist in a variety of forms or habits. Our present concern involves the classification through a fractal dimension based scheme of these hydrometeors, which have themselves been "imaged" by an optical array probe. Although the microphysics of hydrometeor type is a subject in its own right, optical array probe measurements of hydrometeoric mass fluxes and comparisons with weather radar reflectivity data,[1–3] serve to illustrate the importance of this type of study in a broader atmospheric science context.

The nature of the optical array probe data is reviewed in the following section, as are the methods used to analyze it. An efficient algorithm, which allows for the estimation of both the capacity and information dimensions, is the subject of the third section. The fourth section of this paper discusses the application of this algorithm to various synthetic images, and highlights some of the successes and challenges that characterize our ongoing quantification process. Conclusions and future directions are considered in the fifth and final section of the paper.

2. PROBE DATA AND ITS ANALYSIS

For about two-and-a-half decades, Knollenberg-type[4] optical array probes have been used to render *in-situ* digital images of hydrometeors. Such hydrometeors are represented as a two-dimensional matrix, whose individual elements depend on the intensity of transmitted light, as these hydrometeors pass across a linear optical array of photodiodes.[5] Although there are a number of details that need to be considered in processing this probe data,[6] attention is focused here on classification schemes based on hydrometeor type, e.g. plates, stellar crystals, columns, spatial dendrites, capped columns, graupel, and raindrops.[1,7]

In previous efforts,[1,2,5,6] hydrometeor habit has been "quantified" through simple dimensional analyses, and one of these methods[5] also allowed for image enhancement procedures in the analysis of truncated images. More complex approaches have utilized statistical pattern recognition algorithms[7,8] and Fourier-based radial harmonic analysis methods.[9] Despite these efforts, and others not summarized here, in 1985 Heymsfield and Baumgardner[6] concluded that "thus far, no technique has been developed that will classify the more irregularly shaped images that would be representative of rimed particles or spatial dendrites." This, in part, forms the central motivation for the present study.

3. FRACTAL DIMENSION BASED ANALYSIS

3.1 Area-Perimeter Relations

Quite recently,[3] a "roughness" parameter has been employed in the classification of optical array probe data. This "roughness" parameter, α, is formulated in terms of a power-law relationship between the surface area, A, shadowed by the hydrometeor, and its perimeter, P, namely $P \sim A^{\alpha}$. The claim has been made that values of α close to 0.5 were representative of dense, spherical, heavily rimed particles (hail or graupel), whereas values closer to

0.8 were indicative of low-density, irregularly shaped aggregates.[3] Although it is tempting to view α as an estimate of the fractal dimension based on the area-perimeter type scaling rule (see Lovejoy[10]) given above, Marécal[3] advise against this for the following two reasons: the small image size of their probe data (32 by a maximum of 200 pixels), and the possible existence of a non-standard relationship, which could result from variations associated with particle type, between area and characteristic length. Marécal[3] concluded that it was reasonable to consider α as a *heuristic* parameter, that allowed them to distinguish between graupel and snowflake/dendrite populations.

In a spirit that is analogous to that of Marécal,[3] our present goal is to apply a fractal dimension based analysis to similar optical array probe data. Our intention is to perform a quantitative geometrical analysis, based on the capacity and information fractal dimensions, in order to facilitate the possibility of a quantitative classification of hydrometeoric type.

3.2 Capacity Dimension

Following the notation of Leibovitch and Toth,[11] the capacity dimension, d_B, can be written as

$$d_B = \lim_{\epsilon \to 0} \frac{\log N_B(\epsilon)}{\log(1/\epsilon)}, \tag{1}$$

where $N_B(\epsilon)$ represents the minimal number of covering cells (e.g. boxes) of size ϵ required to cover a set, $\mathcal{S}$. Standard box counting suffers from requirements for large data sets, abundant computer memory and lengthy computation time. The box-counting algorithm developed by Leibovitch and Toth[11] requires $N \approx 10^d$ data points, $\approx N d_e$ memory locations and an execution time of order $N \log N$, for an attractor of dimension d, embedded in a space of dimension d_e.

The Leibovitch-Toth algorithm (LTA) proceeds in the "traditional" sense by determining the number of boxes required to make a minimal cover of the set, with the proviso that each box contains at least one element of the set. Box size is then systematically decreased with this minimal-cover operation being executed at each stage. An efficient hashing allows all of the points within a given box to be represented by the same number, then a count of the number of distinct values provides $N_B(\epsilon)$. The slope on a standard Richardson plot[12] of $\log N_B(\epsilon)$ versus $\log(1/\epsilon)$ then provides an estimate of the capacity dimension.

3.3 Information Dimension

Since the capacity dimension does *not* account for the frequency with which the set, $\mathcal{S}$, (whose fractal dimension is being estimated) "visits" the covering cells, the local properties of this set are not distinguishable.[13] Thus the information dimension, d_I, which is used to measure the average information conveyed by knowing the covering cell in which a point of the set $\mathcal{S}$ in question lies in, can be written as

$$d_I = \lim_{\epsilon \to 0} \frac{I(\epsilon)}{\log(1/\epsilon)}. \tag{2}$$

In Eq. (2), $I(\epsilon) = \sum_{i=1}^{N_B(\epsilon)} -P(\epsilon,i)\log[P(\epsilon,i)]$ is an average measure of the information associated with the occurrence of a member of the set $\mathcal{S}$ in a covering cell, while $P(\epsilon,i)$ is the sample probability that a point of the set $\mathcal{S}$ lies in the i^{th} covering cell.[13] The information dimension serves as a lower bound for the capacity dimension,[13,14] i.e. $d_I \leq d_B$, which in addition accounts for the non-uniform distribution of the set $\mathcal{S}$ over the covering cells. The slope on a plot of $I(\epsilon)$ versus $\log(1/\epsilon)$ provides an estimate of the capacity dimension.

4. RESULTS

The capacity and information dimension results, for the simple, two-dimensional binary matrix silhouettes discussed in this section, were obtained through the use of our own implementation of the algorithms described previously. Satisfactory operation of our own code was determined through comparisons with established fractal-dimension estimating software.[15] In this preliminary report, we focus on results gleaned from fractal-dimension characterizations of known, simple synthetic images. It is critical to note that even though we have tested our code on spatially complex images, these low complexity images are more appropriate in the present context of optical array probe data. In particular, capacity and information-dimension estimations as a function of image size for a circle and Koch island are presented.

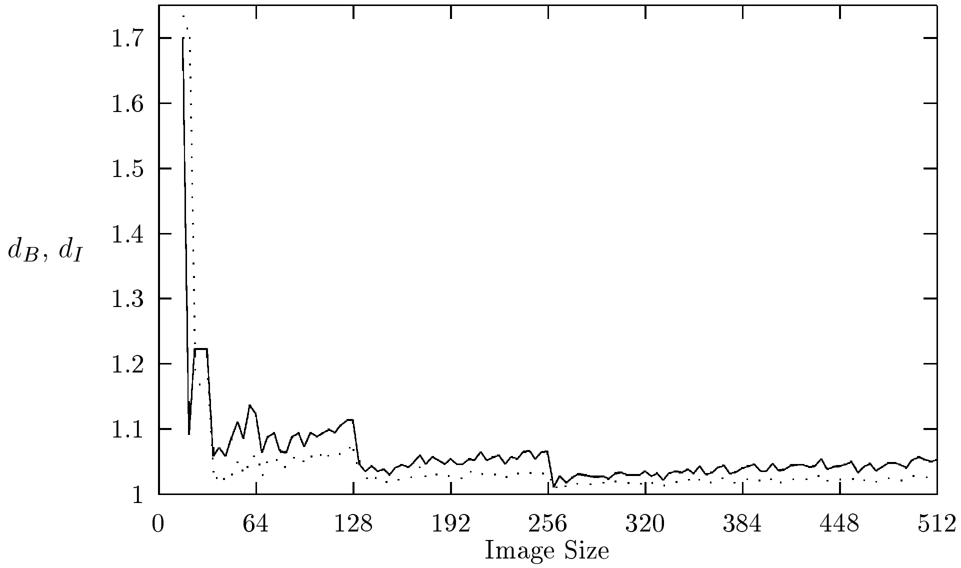

Fig. 1 Estimated capacity (solid curve) and information (points) fractal dimension as a function of image size (circle radius) for a circle.

Perhaps the simplest type of hydrometeor to be imaged, by an optical array probe, would be that of an aspherical raindrop. This particular hydrometeor motivated the fractal dimension calculations illustrated in Fig. 1. In this figure, the capacity (solid curve) and information (points) dimensions are shown as a function of radius for a two-dimensional binary matrix representation of a circle. The initial peak in this data set corresponds to poor fractal-dimension estimations due to the smallness of the data set used. As intuitively expected, these same estimates tend to decrease with increasing data set size, and eventually tend towards the 1.0 theoretical "fractal dimension" for a circle. It is also clear from this same figure that there exists a series of three "plateaus", i.e. regions where both the capacity and information-dimension estimates level off to some more or less constant value. This artificial 'multifractal' effect is associated with the number of points used in the slope calculation described previously. Except for very small images, the information dimension is always smaller than the capacity dimension, as expected from theory.[13,14]

In an effort to move towards an understanding of quite complex hydrometeor habits, a Koch island generator[12] was applied to a square initiator. In the lower part of Fig. 2, capacity (solid curve) and information (points) dimension results are presented as a function of the initiating square's width. As in the case for the circle, poor fractal-dimension estimates are derived for small images; due to the increased complexity of this figure, the "multifractal plateauing" effect is even more pronounced here. In contrast to the results obtained for the circle, and in exception of the theoretical upper-bound stipulation noted previously, the information dimension exceeds the capacity dimension regardless of image size; similar results have been found through the use of other implementations of these same algorithms. Although the theoretical fractal dimension of a "multiple-generation" Koch island is 1.5, the best estimate of this first-generation Koch island lies in the 1.05-1.10 range; since this first-generation Koch island is considerably simpler that the Koch island possessing a 1.5 fractal dimension, this low fractal dimension is not too surprising.

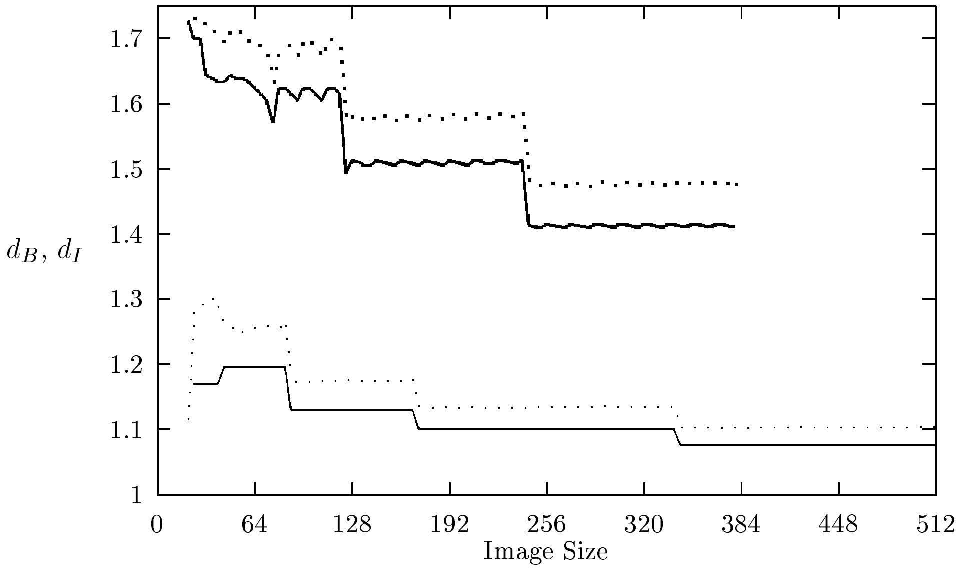

Fig. 2 As in Fig. 1, except that the image size corresponds to the width of the square initiator. Results for the "first-generation" Koch island are presented in the lower part of this figure (light lines and points), whereas the "second-generation" results are illustrated in the upper portion (heavy lines and points).

A more complex Koch island was constructed by a second application of the generator[12] to the first-generation Koch island; the resulting dependence of capacity (dark solid curve) and information (dark points) dimensions as a function of initiating square width is presented in upper part of Fig. 2. Here the results are qualitatively similar to those obtained for the very simple Koch island. It is important to note, however, that the ≈ 1.4 best estimate of the capacity and information dimensions of this more complex Koch island are much closer to the theoretical value mentioned previously.

5. DISCUSSION

Although our conceptualization of nature in terms of fractal geometry dates from at least 1975,[12] the use of this underlying geometrical structure in quantitative image classification is considerably more recent.[3,16] In some senses, the present effort follows the area-perimeter based classification scheme developed by Marécal[3] Since we have also found that image size appears to be the single most limiting factor in a fractal dimension based characterization, we have shown how estimates of the capacity and information dimensions vary as a function of this same parameter, for two extreme habits of synthetic hydrometeors. While it is manifestly clear that estimates of the *absolute* capacity or information dimension are not possible for these small data sets (i.e. 32-64 by ~200 pixels), we have shown here that *relative* fractal-dimension estimates can indeed be obtained. It is thus our expectation that further forward-modeling studies directed at fractal-dimension estimations of the seven hydrometeor types discussed at the outset, will likely lead to a *heuristic* characterization of the various hydrometeors in the same spirit as Marécal.[3] Spectrally deduced estimations of fractal dimension, using the maximum entropy method, are also currently under investigation.

ACKNOWLEDGMENTS

We would like to thank J. W. Strapp for numerous discussions relating to the optical array probes, and T. G. Smith for introducing us to quantitative image classification. We are also grateful to the Natural Sciences and Engineering Research Council (NSERC) of Canada for partial funding of this project.

REFERENCES

1. E. W. Holroyd, III, J. Atmos and Oceanic Tech. **4**, 498 (1987).
2. R. R. Braham, D. A. R. Kristovich, and M. J. Dungey, J. Appl. Meteor. **31**, 237 (1992).
3. V. Marécal, D. Hauser, and C. Duroure, Atmos. Res. **29**, 179 (1993).
4. R. G. Knollenberg, J. Appl. Met. **9**, 86 (1970).
5. A. J. Heymfield and J. L. Parrish, J. Appl. Meteorol. **17**, 1566 (1978).
6. A. J. Heymfield and D. Baumgardner, Bull. Amer. Meteor. Soc. **66**, 437 (1985).
7. M. M. Rahman, E. A. Quincy, R. G. Jacquot, and M. J. Magee, J. Appl. Meteor. **20**, 521 (1981).
8. M. M. Rahman, R. G. Jacquot, E. A. Quincy, and R. E. Stewart, J. Appl. Meteor. **20**, 536 (1981).
9. C. Duroure, J. Rech. Atmos. **6**, 71 (1982).
10. S. Lovejoy, Science **216**, 185 (1982).
11. L. S. Leibovitch and T. Toth, Phys. Lett. A **141**, 386 (1989).
12. B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman and Company, 1982.
13. J. J. Sarraille, Inaugural Conference for a Society for Chaos Theory in Psychology, San Francisco, California, August 15-16 (1991).
14. H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*, Springer-Verlag, 1992.
15. J. J. Sarraille, and P. DiFalco, *FD3* (fractal dimension estimation software available free from `lyapunov.ucsd.edu` via anonymous ftp in directory `/pub/cal-state-stan`, 1991).
16. T. G. Smith Jr., W. B. Marks, G. D. Lange, W. H. Sheriff Jr., and E. A. Neale, J. Neurosci. Methods **27**, 173 (1989).