EUROGRAPHICS - IEEE VGTC Symposium on Visualization (2005)
K. W. Brodlie, D. J. Duke, K. I. Joy (Editors)

Pointillist and Glyph-based Visualization of Nanoparticles in Formation

P. Coleman Saunders[1] Victoria Interrante[1] S.C.Garrick[2]

[1]Department of Computer and Information Science,
[2]Department of Mechanical Engineering
University of Minnesota, Minneapolis, Minnesota, USA

Abstract

In this paper we offer new, texture-based methods for the visualization of multivariate data. These methods aim to more effectively convey the results of calculations simulating the formation of nanoparticles in turbulent flows. In these simulations, an entire distribution of nanoparticles is computed at every point across a two-dimensional slice of the data space, for every time step. Previous visualization methods have relied on multiple separate images to convey summary statistics about the datasets, including mean diameter and standard deviation of particle sizes. We introduce new methods based on texture which aim to enable the integrated understanding of the entire distribution of values at each point across the domain in terms of both summary statistics at each point and particle counts for various sizes of particles. Pointillism is used to represent the data at each point across the data range as a high-resolution texture. Circular glyphs can also be used to form a more discrete, spot-based texture, in which different characteristics of the distribution are encoded in various features of the spots.

Categories and Subject Descriptors: I.3.3 [Computer Graphics]: Picture/Image Generation

1. Introduction

Vapor-phase methods are of primary importance in active research into the production of nanoparticles. Simulating these production processes may allow us to optimize and thereby increase the efficiency of the nanoparticle production process [MG03], facilitating economical production of materials with unique and useful properties. Use of new methods for direct numerical simulation of coagulating aerosols have produced multi-dimensional datasets which capture the spatial and size distribution of nanoparticles at various time intervals in the formation process [Giv89]. Visualization of the geometric properties of these nanoparticles represents an interesting and relevant problem for scientists because direct visualization of physical nanoparticles is impossible- no camera device exists capable of capturing images with sufficient speed to give a snapshot at small time intervals. Additionally, the nanoparticles possess the particularly vexing quality of scattering light consistent with their diameter raised to the sixth power; therefore, were it possible to image the particles directly, scalar considerations would either leave the smallest particles invisible or produce an intractably large image.

We have examined computed datasets which describe quantities of nanoparticles in formation at every point in highly intermittent turbulent flows. This data is separated into individual slices of counts of particles one, two, four, eight, sixteen, and thirty-two nanometers in size. First, we consider visualization of summary statistics of the dataset, consisting of the mean diameter and standard deviation of particle sizes at each point. We furthermore apply our methods to the particle quantity slices with the goal of visualizing all size distributions simultaneously.

2. Summary statistics visualization

Researchers compute the mean diameter and standard deviation of particle sizes for every point across the range of data seeking a variety of insights- *e.g.* the deviation at a location can inform scientists about the strength

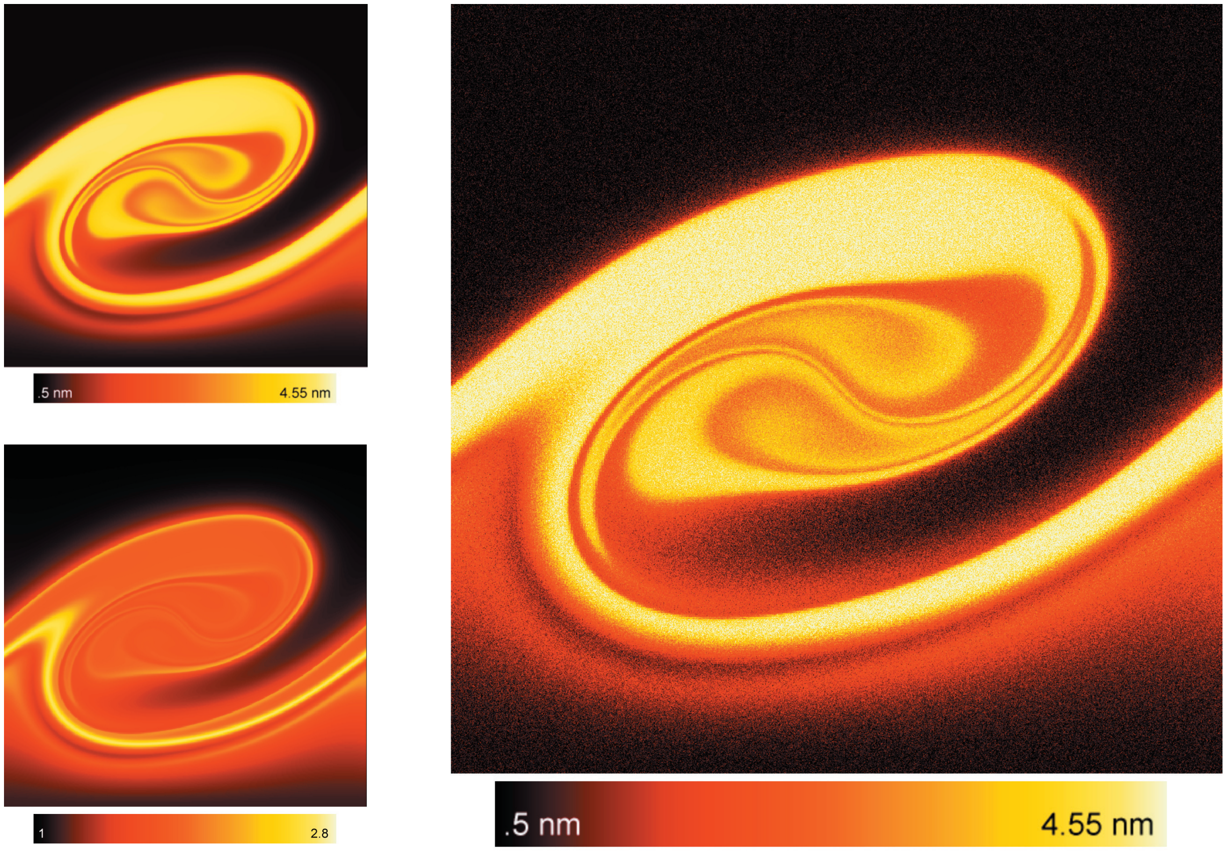


Figure 1: *Direct renderings of mean diameter and standard deviation versus pointillist rendering on the heated-object scale.*

of the effects of transport on the nanoparticles [MG03]. Previous visualization methods for summary statistics have simultaneously utilized two separate images with different contexts. One image visualized the mean diameter of particles at each point and while another conveyed the standard deviation. In Modem and Garrick's work with this dataset, both direct visualizations are mapped onto a rainbow color scale [ibid].

There can be problems with the rainbow color scale- even with a legend the ordering of colors is not necessarily intuitive, and it is not perceptually linear from a luminance perspective [RT98] which can give rise to visual artifacts. Research has shown that simultaneous mental combination of spatially-referenced features in multiple images is at best difficult and at worst intractable [*e.g.* TV98]. We have rendered separate-context images (figure 1) in the style of Modem and Garrick's rainbow images [MG03] for comparative purposes, albeit using a heated-object scale to avoid color-related artifacts. In the separate renderings of mean diameter and standard deviation (figure 1 left), each statistic is explicitly represented throughout; however, it is not trivial to relate a given point in one rendering to its equivalent point in the other. Our goal is to combine these two renderings into a single, intuitive image.

2.1 Pointillism and related work for summary statistics

The concept of applying painterly effects to data visualization in computer graphics has rich historic precedent. Laidlaw *et al.* [LAK*98] and Kirby *et al.* [KML99] offered several approaches to artistic rendering of scientific data using layering and "paint stroke" glyphs. Healey *et al.* [HTE*04] have investigated perceptual principles for effectively conveying data using nonphotorealistic visualizations. There is a wide range of more artistically-motivated algorithmic treatments of painterly effects for the purposes of computer-enhancement of the artistic process or exploration of computer graphics potential [HAE90][GG01]. Pointillism was suggested as a useful method for visualizing cartographic data as early as 1953 [JEN53] and we feel it is a natural vehicle for our datasets as well. Indeed, the term pointillism has already been mentioned in correlation with nanoparticles in formation [RT98].

We use Pointillism, a technique through which we represent a single mean and standard deviation pair with a larger area square of pixellated texture in the output image. Pointillism allows us to combine the mean diameter and standard deviation statistics in a single image by mapping the average color in an area to the mean diameter, and the amount of color-wise variation in the texture pattern to the standard deviation. For each pixel in a statistic pair's texture-square, we generate random normally distributed data values from the distribution defined by the underlying statistics. We have utilized the Box-Muller Polar transform for generating normally distributed numbers from a pseudorandom uniform distribution [BM58].

In our figures, the random normal values are then mapped to the conventional heated-object colorscale before output. The HOS was chosen to exploit human sensitivity to luminance in the yellow-orange hue [LH92] and for its familiarity to the chemical engineering community. The resultant image (figure 1 right) shows both statistics by mapping the mean to the average color value in an area, and the standard deviation to the diversity or "graininess" of color within that area. Viewed at the original scale, this image gives an intuitive impression of the overall mean particle dimensions as or more effectively than the separate images (figure 1 left). To get a feel for the standard deviation in an area, we need only zoom for a closer view of the graininess pattern in an area.

We believe this type of visualization is quite effective for showing us these statistics for particles in a given area; however, it may still lack some the intuitive qualities we strive for. One issue is that the user must interpret particle size from color intensity. The other is that the standard deviation of the distribution is not very obvious at the farthest scales, requiring either that the image be viewed on a very large, high-resolution display or that the user interactly zoom the image to achieve the insight s/he desires.

An alternative approach is to portray the mean diameter statistic more directly. A second method we have developed generates spot glyphs of which the perceptual sizes are representative of the nanoparticle sizes found in the corresponding regions of the flow.

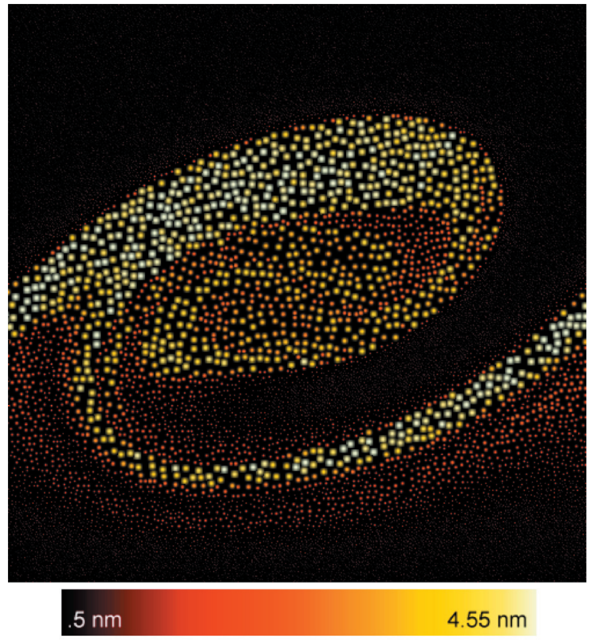


Figure 2: *Spot glyph rendering of mean diameter and standard deviation on the heated-object scale.*

2.2 Spot glyph production for summary statistics

The perceptual size of our spots is mapped to the mean diameter at a point. Our method uses Gaussian shading to control the perceptual size for a given spot. To show the standard deviation in the data, we follow a model akin to our pointillist techniques. We want to vary the diversity of spot sizes in a given area in relation to underlying standard deviation statistics, *i.e.* in a texture-square representing a high standard deviation we would expect to see spots with a larger variety of sizes than in a texture-square representing a low standard deviation, with the average size of spots in a square linked to the underlying mean diameter statistic.

To this end, for each potential spot we calculate a value Σ, which equals the mean diameter μ plus or minus a normally distributed stochastic variable R multiplied by the underlying standard deviation σ. The Σ value is used in the calculation of the Gaussian shading function for a given spot. In other words, the Gaussian shading (and hence perceptual size) of a given spot varies generally with the mean diameter statistic, modified by a random normal distribution around that mean defined by the standard deviation. This mapping is quite intuitive- it requires no additional information such as keys or a legend to convey variation over space, and allows us to see not only the mean diameters across the entire space, but also the distribution of the standard deviation statistic in a given area. Thus we can visualize not only the diameter distribution through use of the range of spot sizes, but also the diameters' standard deviation distribution in an area through the varying of diversity of spot sizes in that area. Visualizing both the distribution of particles at each point (the summary statistics) and the distribution of these statistics across the whole space is a novel contribution of these methods.

2.3 Glyphs and spatial frequency for summary statistics

One problem in using spot glyphs of varying size to convey data is correctly handling the negative space. If the layout of spots is strictly uniform, smaller spots (having smaller footprints) will appear further apart than their larger cousins, which fill the negative space more fully. Healey and Enns have shown glyph density to be an effectively detectable artifact [HE99]; therefore, when we generate the spots we must take care to maintain a perceptually similar spatial frequency for small spots and large spots. We should also avoid implying false information about the density of the underlying nanoparticle distribution by packing smaller spots more tightly and hence maintaining a consistent spatial frequency across both positive and negative space of the pattern. This method also avoids potential misinterpretations of the data due to regularity of glyph spacing, which is another easily detectible artifact [*ibid*].

A conventional Poisson method of distributing spots would read like this- pick a random location on the output field and calculate the size of a spot at that point based on the mean diameter and standard deviation data. If that spot conflicts or overlaps with an existing spot, throw it away and repeat until the image buffer is full, otherwise, place the spot and continue until the buffer is full. There are two problems with this algorithm- the random nature of picking spot locations implies that the rate of filling the image decreases as the image becomes more full and more spot-placement conflicts occur. The other issue is that there is no efficient method of determining when the algorithm has filled the space with a "sufficient" number of spots; that is, there isn't a fast way to determine when the packing is complete.

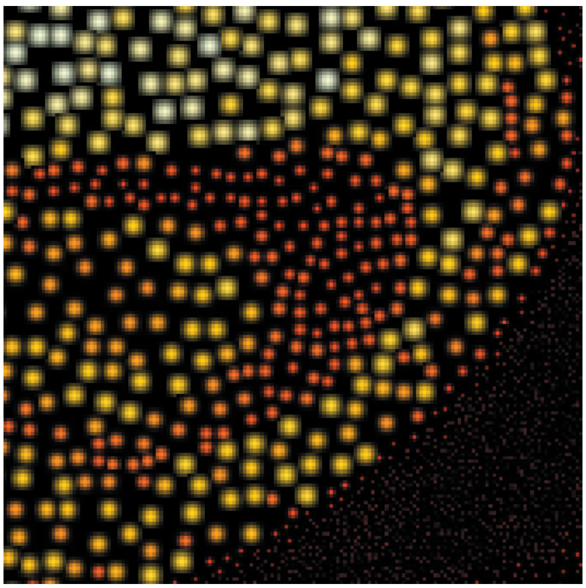


Figure 3: *Enlargement of spot glyphs shows size and packing detail*

Secord presents a novel method of evenly packing spots through the use of weighted centroidal Voronoi diagrams [Sec02]. This method is effective but costly. We have reasonably approximated a full Poisson packing through more deterministic methods. First, we generate a regular grid of potential locations to place the center of a glyph. The density of this grid must be sufficient to maintain a high spatial frequency for the smallest possible spot size, a single pixel. We then jitter the prospective spot locations by a random small amount in any direction to produce a non-regular but *practically* constant density distribution of potential spot locations throughout the output field.

For each potential spot location, we calculate the footprint of the would-be spot. By consulting the output buffer so far, we can determine whether the footprint overlaps with an existing spot. In the overlap case, we simply throw the spot away and move on to the next potential location, otherwise we place the spot and move on. The high spatial frequency of the underlying jittered grid suggests that in most cases spots will appear in a well-packed manner without covering each other unless we so desire. For applications in which spot density is a concern (*e.g.* several layers of spots on top of each other), we can modify the overall spot density by changing the density of the underlying grid of potential spot locations. This stochastic method cannot guarantee an optimal packing, but it is computationally efficient and avoids most serious visual artifacts due regularity or spatial frequency.

3. Particle count data visualization and color

Following our work with a summary statistics describing mean diameter and standard deviation of particles over space, we turned our attention to the underlying data. In this data are six cospatial slices describing the quantity of particles at each point for six sizes of particles: one, two, four, eight, sixteen, and thirty-two nanometers in diameter. Researchers want to examine these datasets simultaneously to learn about the diversity of types of nanoparticles (different size-bins, that is), the relationships between these sizes, and their quantities in an area. The data are used in various ways. Particle size affects chemical reactivity on the surface of the particles; hence, one visualization of particular interest will convey information about the homogeneity (relative diversity) of particle sizes by particle counts at a point. Visualization of the particle counts also assists scientists in describing their behaviour, *e.g.* revealing a deficit of large particles in an area demonstrates that particles grow more slowly in that region [MG03]. To the end of supporting such visualizations, we have extended our pointillist and glyph techniques.

A wide variety of research supports the efficacy of using color as a differentiating variable among data. (*e.g.* [LH92][RK01]) In our case, values from a given size-bin (*e.g.* all particles 4 nm. in diameter) are consistently tied to a single colorscale to differentiate respective quantities from those of larger or smaller-sized neighbors. Our choice of colors and transformation (mapping) functions is not trivial, however, as our display methods and the datasets themselves present some challenges. We believe the most effective visualization uses a unique transfer function between data value and color for each size-bin of particles and employs perceptually equiluminant colorscales (figure 4).

We used unique transfers functions for each size of nanoparticles because of scale considerations- the datasets occupy a high dynamic range. The maximum number of particles of size 8, for example, is many orders of magnitude greater than the maximum number of particles of size 32. Using distinct transfer functions for each color/size avoids problems of data loss that can result from mapping vastly different ranges with the same function, *e.g.* if we use the same function for 32 nm. particles as 8 nm. particles, all colors for 32 nm particles will appear only in the lower part of the colorspace because the maximum quantity of size 8 particles is so much greater than that of size 32 particles.

We chose to link our colorscales to the particle counts directly and not to the percentage of all particles at a point that a given size represents. Using percentages sacrifices all information about the actual number of particles- *e.g.* twenty percent of one hundred particles is significantly different from twenty percent of one million particles, but a percentage-based method would show them as the same color, possibly leading to erroneous insight into the data.

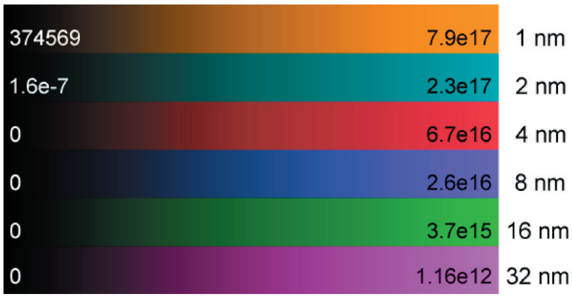


Figure 4: *Perceptually equiluminant colorscales for use with all renderings of multiple sizes and counts of nanoparticles.*

Because of the close proximity of differing colors in our representations, we utilize perceptually equiluminant colorscales. These colorscales ensure that two colors representing the same proportionate value on two colorscales appear equally bright. While *perceptually equiluminant* is, at best, an approximation across different viewers, methods have been developed to facilitate the selection and evaluation of potential colorscales for perceptual equiluminance. One such system is Kindlemann's Face-based Luminance software [KRC02]. In applying both the pointillist and glyph techniques to our datasets we have used perceptually equiluminant colorscales which represent the best possible balance (in our estimation) between diversity of color and equiluminance, as shown in figure 4. It should be noted that factors outside the scope of this work can also affect perceptual equiluminance, such as monitor calibration and CMYK color printing.

3.1 Pointillism and motivation for particle count data

To make use of the pointillist technique with these datasets, we have modified the stochastic process first used with summary statistic datasets to show standard deviation over an area by variation ("graininess") in the texture. As we have no explicit deviation data in the second group of datasets, we can exploit stochastic sampling to convey data across all datasets. For each pixel, we choose randomly which size-bin to visualize in the final output image and subsequently map that data value to its corresponding colorscale. This is a simple process for combining the multiple datasets and may initially seem rather naïve; however, to justify the utility of the method we need only compare it to alternatives that may fail to provide similar insight.

First we examine the notion of random dataset selection in the sampling process. Critics might suggest that choosing but one data value to visualize for each pixel is unnecessary, and we might instead perform blending of all datasets to see everything at once. The following visualizations illustrate how such a method breaks down: figure 5 illustrates direct visualizations of all particle sizes for comparison purposes. Figure 6 compares the blending of direct renderings of all nanoparticle sizes versus a pointillist version. It is apparent that the blending method fails to give discrete impressions of each data subset while the pointillist method retains distinctness, most obviously in areas of high particle count, but also those of very low particle count.

Instead of randomly choosing which size to render for a given pixel in the output texture, we might also consider a fully regular spacing for the rendering of each data subset. For example, we can consider the case of a square glyph of pixels in which the upper-left pixel always corresponds to particles of size 1, the upper-right to size 2, and so on. Various work discusses the problematic nature of regular sampling across a spectrum of graphics research (*e.g.* [M87][C86]). Comparisons with a picture of the same data rendered with our random methods shows the superiority of the random-choice method- even if we zoom in very close (see figures 7, 8) the colors in the stochastic pointillist rendering are far more differentiable than the ordered version. We argue that regularity in the texture substantially interferes with our ability to discern small color variations across space.

3.2 Glyphs for cospatial particle count datasets

Our spot glyphs have been adapted to give a different view on these datasets. Instead of a single color representative of the mean diameter of particles at a point, we now present a target motif glyph using concentric rings of color tied to underlying particle count values. The order of colors/sizes is progressive; that is, the inner-most ring color corresponds to the number of particles of size 1 at the center of the target, the next ring out corresponds to particles of size 2, and so on. Again we have utilized the perceptually equiluminant colorscales (figure 4) to ensure that one colorscale does not gain an unfair "perceptual advantage" over any other. The glyph-packing algorithm is the same as that for the spot glyphs used with summary statistics, detailed prior in section 2.3.

While the round target glyph is a natural extension to the spots we utilized for summary statistic visualizations, it is not the only glyph we considered- we also examined line glyphs with colors arranged in a pattern, and square targets. These options have failed to produce as robust a visualization. Interference due to the perceived orientation of the line glyphs and the linear qualities of the edges obfuscated the relative continuity of the underlying data. The same effect can be seen with the square targets due their linear edges, though it is not as pronounced. In addition to these factors, the rounded qualities of the target glyph more intuitively equate to scientists' mental representations of the nanoparticles.

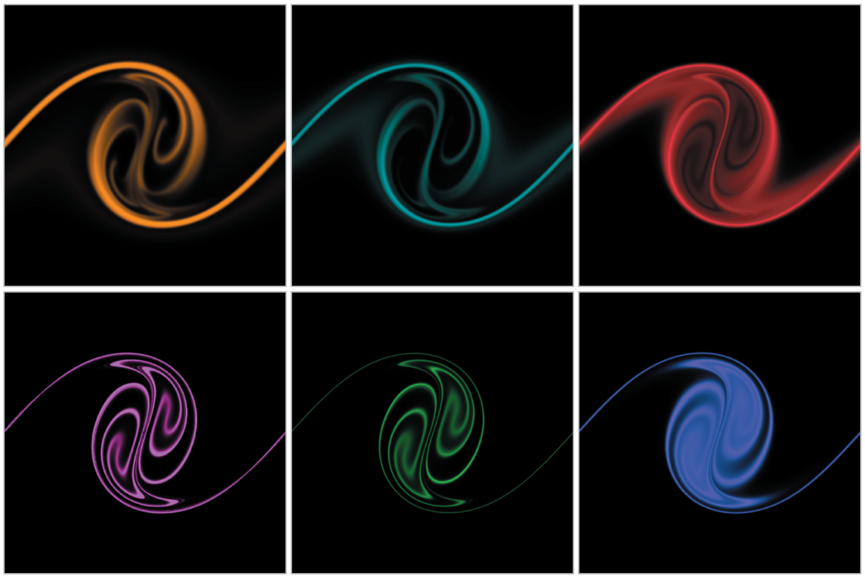
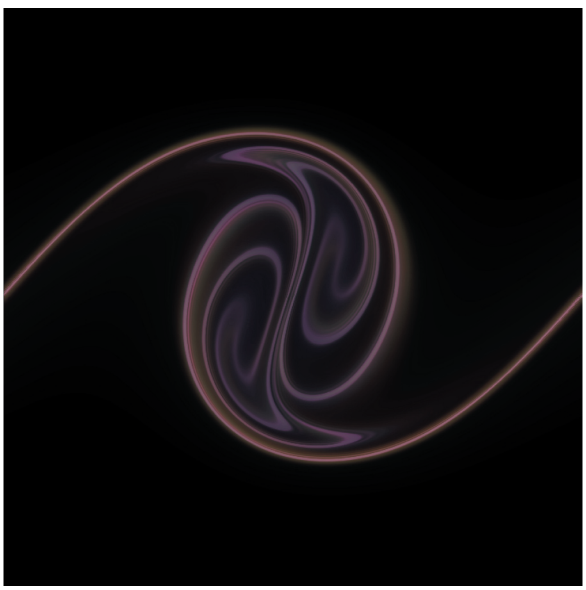


Figure 5: *Direct renderings of nanoparticle counts at each point for six sizes of nanoparticles. Clockwise from the upperleft: counts for 1, 2, 4, 8, 16, and 32 nm. sizes. Colorscales follow from figure 4.*



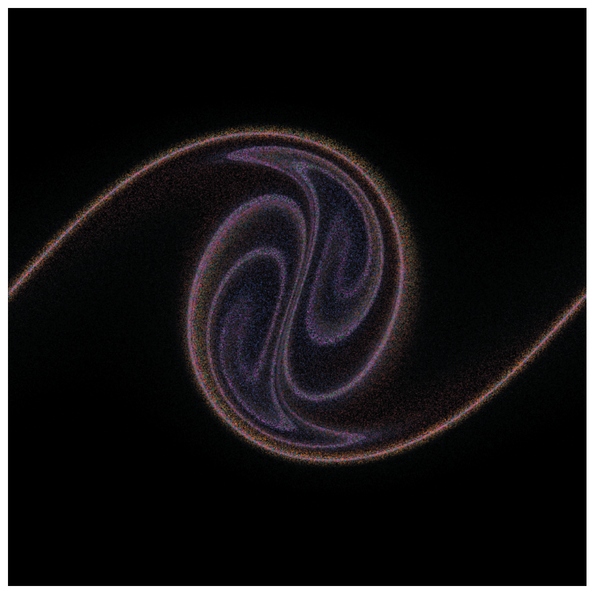


Figure 6: *Blending versus pointillism for showing multiple datasets. At left, blended rendering for all sizes of nanoparticles. At right, pointillized rendering for all sizes preserves relevant details.*

The target representation may lead to additional insights about the data which are not immediately obvious from the pointillist visualization. The glyphs have the ability to visually emphasize borders in regions of the data where numbers of particles change quickly. For example, if we examine the target rendering in figure 9, we can gain simultaneous insight to homogeneity and relative size of particles in the flow eddy. The brighter, multi-hued targets at edges of the interfaces in the flow pattern in comparison with the smaller, consistently blue targets in the core of the eddy may imply that particles are coagulating more rapidly at these edges than in the core [MG03].

4. Conclusion

We have offered visualization methods to effectively convey information about the formation of nanoparticles in computed two-dimensional slices of multidimensional incompressible particle-laden flows. We have applied our techniques to the visualization of summary statistics (mean diameter and standard deviation) for nanoparticle sizes through use of a pointillist texturization technique, in which the average color over space represents the mean diameter and the variegation ("graininess") represents the standard deviation. We have also derived methods for the production of spot glyphs intuitively representative of the

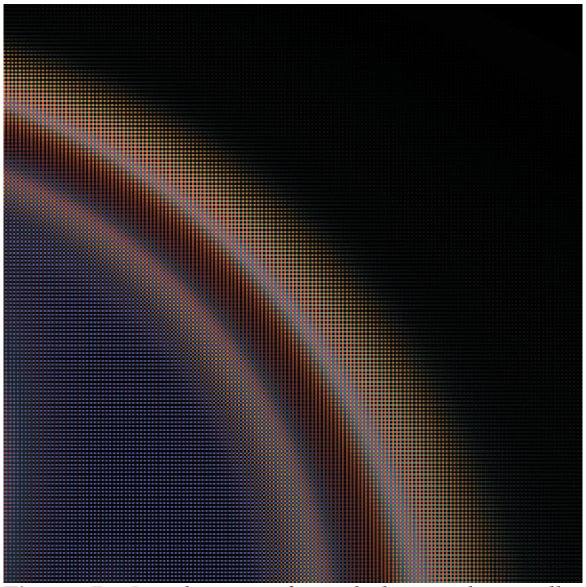


Figure 7: *Detail view of regularly-spaced pointillist rendering.*

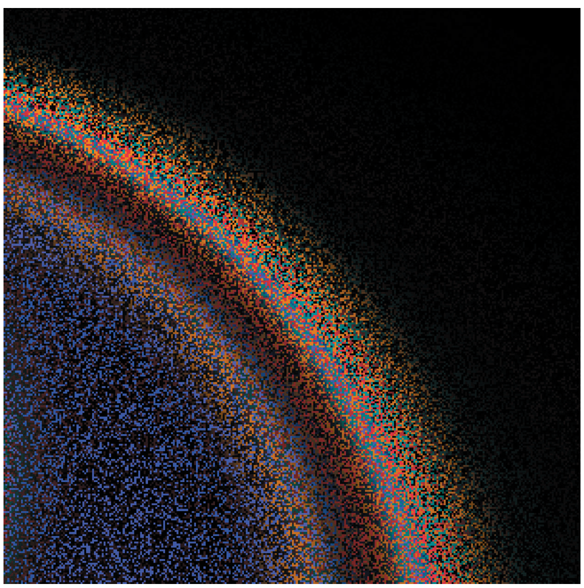


Figure 8: *Detail view of stochastic pointillist rendering maintains the integrity of the visualization.*

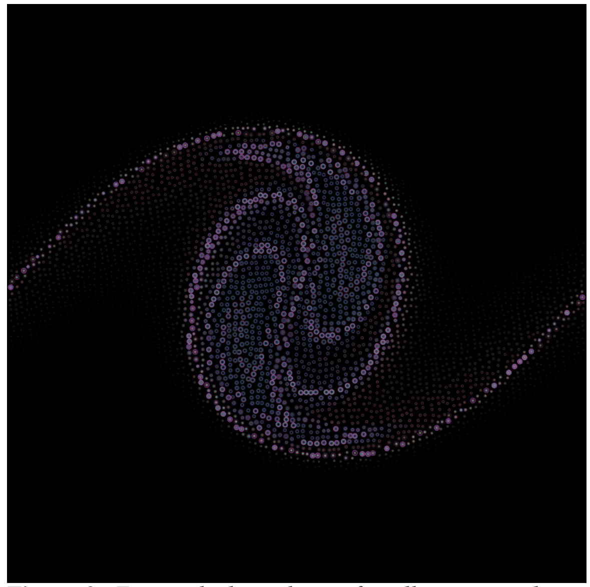


Figure 9: *Target glyph rendering for all nanoparticles in formation.*

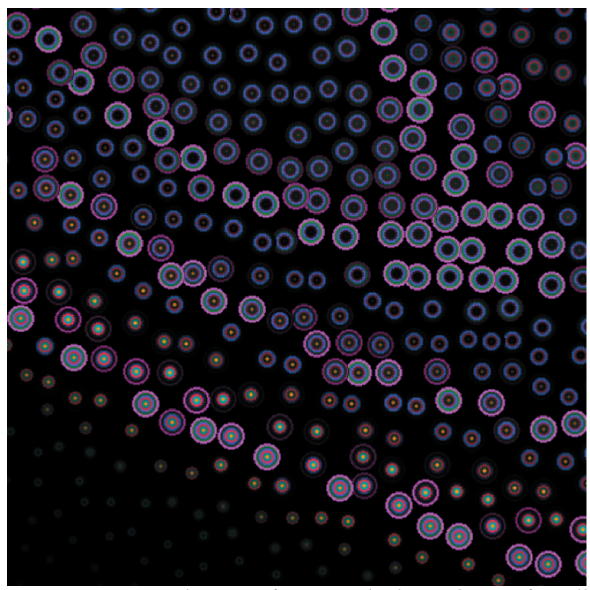


Figure 10: *Detail view of target glyph rendering for all nanoparticles in formation.*

summary statistics dataset in terms of the spots' diameter and variation over space. Furthermore, we have extended both methods for application to the underlying particle count data in which we are interested in the quantities of various sizes of nanoparticles forming at each given location. The pointillist method here allows us to simultaneously visualize an arbitrary number of distributions through stochastic sampling and equiluminant colorscales. The target glyph has the potential to represent a great deal of data in underlying distributions in a unique, compact form. In our case, these techniques have been applied to research scientists' data of nanoparticles in formation in turbulent flows; however, we believe the simplicity of the algorithms, ease of implementation and robustness of the techniques should allow their use in a wide variety of situations in which simultaneous visualization of multiple distributions is called for.

Acknowledgments

This work is supported by a grant from the Army High Performance Research Computing Center (AHPCRC) under the auspices of the Department of the Army, Army Research Laboratory (ARL), under Cooperative Agreement number DAAD19-01-2-0014., with additional funding from NSF (CTS-0324898). The content does not necessarily reflect the position or the policy of the government and no official

endorsement should be inferred. Computing resources for the data calculation were provided by the Minnesota Supercomputing Institute and the Digital Technology Center.

References

[BM58] Box, G.E.P., Muller, M.E.: A Note on the Generation of Random Normal Deviates. *Ann. Math. Stat. 29*, (1958), 610-611.

[C86] Cook, R.: Stochastic Sampling in Computer Graphics. *ACM Transactions on Graphics, 5,* 1, (1986), 51-72.

[Giv89] Givi, P.: Model Free Simulations of Turbulent Reactive Flow. *Progress in Energy Combustion Sciences, 15*, (1989), 1-107.

[GG01] Gooch, B., Gooch, A.: Non-Photorealistic Rendering. A. K. Peters Ltd, (June 2001).

[GCS02] Gooch, B., Coombe, G., Shirley, P.: Artistic Vision: painterly rendering using computer vision techniques. *Proceedings of the second international symposium on Non-photorealistic animation and rendering*, (2002) 83 - ff.

[Hae90] Haeberli, P.E.: Paint by numbers: Abstract image representations. *Computer Graphics. In Proc. SIGGRAPH '90, 24*, (Aug 1990), 207-214.

[HE99] Healey, C., Enns, J.: Large Datasets at a Glance: Combining Textures and Colors in Scientific Visualization. *IEEE Transactions on Visualization and Computer Graphics 5*, 2 (April-June 1999).

[HTE*04] Healey, C., Tateosian, L., Enns, J.T., Remple, M.: Perceptually Based Brush Strokes for Nonphotorealistic Visualization. *ACM Transactions on Graphics*, 23 , 1, (Jan. 2004), 64-96.

[Jen53] Jenks, G.F.: Pointillism as a cartographic technique, *Professional Geographer, 5,* 5, (1953).

[KGC00] Kaplan, M., Gooch, B., Cohen, E.: Interactive Artistic Rendering. *Proceedings of the first international symposium on Non-photorealistic animation and rendering*. (2000), 67 - 74.

[KRC02] Kindlmann, G., Reinhard, E., Creem, S.: Face-based Luminance Marching for Perceptual Colormap Generation. *Proceedings of IEEE Visualization*, (2002), 299-306.

[KML99] Kirby, R.M., Marmanis, H., Laidlaw, D.H.: Visualizing Multivalued Data from 2D Incompressible Flows Using Concepts from Painting. *Proceedings of IEEE Visualization*, (1999), 333-340.

[LAK*98] Laidlaw, D.H., Ahrens, E.T., Kremers, D., Avalos, M.J., Readhead, C., Jacob, R.E.: Visualizing Diffusion Tensor Images of the Mouse Spinal Cord. *Proceedings of IEEE Visualization*, (1998), 127-134.

[LH92] Levkowitz, H., Herman, G.: Design and Evaluation of Color Scales for Image Data, *IEEE Computer Graphics and Applications, 12*, 1, (Jan. 1992), 72-80.

[M87] Mitchell, D.: Generating Antialiased Images at Low Sampling Densities. *ACM Computer Graphics, 21,* 4, (1987), 65-72.

[MG03] Modem, S., Garrick, S.C.: Nanoparticle Coagulation in a Temporal Mixing Layer Mean and Size-selected Images. *Journal of Visualization, 6*, 3, (2003), 293-302.

[RS03] Riebeek, H., Moore, S.: Building Tunable Matter. *IEEE Spectrum Online*. http://www.spectrum.ieee.org/WEBONLY/wonews/jul03/nano.html. (July 2003).

[RT98] Rogowitz, B., Treinish L.A.: Data Visualization: the End of the Rainbow. *IEEE Spectrum*, (Dec. 1998).

[RK01] Rogowitz, B., Kalvin, A.: The "Which Blair" Project": a quick method for evaluating perceptual color maps. *Proc. of IEEE Visualization*, (2001), 183-190.

[Sec02] Secord, A.: Weighted Voronoi Stippling. *NPAR 2002: Symposium on Non-Photorealistic Animation and Rendering*, (2002), 37-43.

[TV98] Tomat, L., Vicario, G.B.: Simultaneous Comparison of Vertical Line Lengths , *Perception ECVP 98 Supplement*. (1998)

[WK95] Ware, C., Knight, W.: Using Visual Texture for Information Display. *ACM Transactions on Graphics 15*, 1, (Jan. 1995), 3-20.