# Chapter Five: Future Work

There are many open questions regarding the DDS data visualization technique. This chapter outlines four lines of inquiry.

1) Augmenting the DDS visualization toolkit with knowledge-based systems to aid users in creating visualizations of their data with DDS. The DDS visualization toolkit enables users to interactively create DDS visualizations of multiple layers of data. The user can interactively change each of the display parameters associated with a layer: the color, size and density of the spots, the layer order, and animation parameters such as the speed and direction of spot motion.

2) Extending the DDS visualization technique to higher-dimensional data, such as volume visualization or visualization of data on a 3-dimensional surface, such as a geographical terrain.

3) Investigating further the cognitive and perceptual issues surrounding DDS. How do people interact with the DDS visualization system and what impact does real-time manipulation of the different parameter mappings have on understanding? This is an interesting area from both a human-computer interaction viewpoint and a perceptual learning stance.

4) Finally, animating DDS layers, which produces such a dramatic impact, is the most interesting and practical direction for future work. I discuss animation first.

## Investigating Perception

### *Visual Perception of Animated DDS Layers*

Investigating the effect of animating a DDS layer on its visual salience is not only the logical next step in the development and evaluation of DDS but also the most promising. Although the DDS alpha-blended layers are robust in the presence of distractors, animation increases the visual salience of a layer so dramatically that it is tempting to expect that animation can significantly increase the maximum number of layers that can be discriminated with DDS, as well as improve the understanding that results.

Several initial questions regarding DDS animation are described below. Because it is impossible to show the power of animated DDS layers with print medium, some examples of pre-specified animations are provided on the CD included with this dissertation.

How much does animation of a DDS layer increase its visual salience? The intuitive answer, backed by many observations, is that animation greatly increases visual salience of a DDS layer. Picking out an object moving differently against a background of other moving objects has clear survival benefits. Motion increases visual pop-out of an object and draws our attention to the object. Because the spots in a layer all move together as a unit, with the same direction and speed of motion, animation increases the perception that the spots form a coherent group. Movie 1 illustrates these effects with uniform fields of spots.

It would be interesting to explore how much animation increases visual salience of a layer through an experimental evaluation similar to the one described in Chapter Three. I believe that if the target layers are animated, performance on the overlap estimate and intersection sketch tasks would be even less affected by the presence of distractors, and perhaps the number of possible distractors could be increased significantly. Movie 2 presents an example of animated target layers from the *Color-Color* session in the main study. Compared to the static image, animated targets are much more salient.

Does animation change how many levels of transparency we can discriminate and therefore how many different levels of data we can perceive in a DDS layer? While viewing DDS images of the SEM data, several people have remarked that the animation drew their attention to areas in the images where the data values were low in magnitude, which they had previously missed in the static images. Movie 3 illustrates this with the climate data – during the month of August the Rocky Mountains in Colorado have measurable levels of ground-frost – this may go unnoticed in the static image, but once the ground-frost layer is set in motion the data stands out clearly.

Animation attracts our attention to values with low magnitude that may go unnoticed in the static image. Low data values produce faint spots – once the spots are set in motion they suddenly become more noticeable. This does not mean that we could not see the faint spots before animation, just that they require less effort to see when moving. It is interesting to note that, once the faint spots are made more salient with animation, the increased salience persists momentarily even when the animation stops.

Does animation improve our ability to see fine spatial details in the data, i.e. does it increase the perceived spatial resolution of DDS? Again, observations of animated DDS images indicate the answer is yes. In this example, animation revealed a small, isolated patch of data that was either not sampled due to the spacing of the Gaussian spots, or just not noticed due to its size and isolation. Movie 4 presents the bat habitat visualization shown in Chapter Two with the habitat range of the western red bat animated. There is a small, isolated area where the western red bat has been sighted in the North-west near Seattle. The area is small enough that it falls between spots and is not sampled in the static image. Once the layer is animated, the area is alternately sampled and unsampled, which causes a flickering effect of on/off. This is very noticeable; in fact, the isolated area was not noticed until the layer was animated.

Animation also helps reveal details in boundaries of the data. This effect is especially strong for large spots, within which the fine details of a boundary are well displayed, but between which perceptual filling-in of the boundary does not work well because the distance between large spots is too great. Animation of large spots across a finely detailed boundary produces remarkable results – the impression it creates is that the entire scene is visible through large "holes" or windows. Instead of the data being perceived as hidden or missing, it is as if one only has to shift one's head slightly to see everything there is to see. Movie 5 illustrates this with a SEM data layer.

Consider the example with large spots shown in Movie 5. If boundary information is revealed over time through animation, how long does the perception persist once the animation stops? How long must a viewer watch the animation to create an accurate mental model of the data that he or she can remember?

What role does attention play when watching an animation? Does attention influence how many layers can be animated at once without causing visual interference, or without confusing or distracting the viewer? One of the powerful characteristics of visually discriminable DDS layers is that the viewer can attend to one or two arbitrarily chosen layers, temporarily ignoring others; the experimental results presented in Chapter Three show this conclusively. In a static image, attention shifts are based on color, when animation is added to certain layers, attention shifts can be based on the speed and/or direction of motion. If more than one layer is moving with the same speed and direction, color again becomes the discriminating factor. Movie 6 shows the SEM visualization from Chapter Two, with two layers animated. This animation illustrates how animation of two or more layers improves the perception of not just the animated layers, but of the static layers as well.

How well does animation help discriminate among spots with similar colors and sizes? Can multiple layers with the same size and color be discriminated based on speed and/or direction of motion alone? Movie 7 shows an example with the bat habitat ranges.

Animated DDS layers can produce the perception of depth or of volume – how can this guide future data visualizations? The perception of depth occurs in a display with several layers with the same color and direction of motion, but with different sizes and speeds of motion. Movie 8 illustrates this effect – the result is like watching snow falling outside on a winter night.

Does spot size influence how effective an animation can be? Do some colors animate better than others? Does an animation ever interfere, or mask the static layers?

In addition to the linear, uniform-velocity animation shown in the movies provided, what other forms of motion are effective? Rocking, scaling, speed that is data-dependent, random motion, coherent flock-like motion: all are interesting areas of exploration.

### *Perceptual Learning and Interaction with the DDS Visualization Toolkit*

What does exploration with the DDS visualization system teach the user about the data he or she is looking at?

Does interaction produce better understanding? Rheingans [1993] showed in her dissertation that interaction significantly increases a participant's understanding of data. It would be interesting to repeat the study presented in Chapter Three, with the participants able to choose the display mappings for the target and distractor layers interactively. I would predict that performance would be better yet if this were the case.

I have found interactive control of the display parameters to be invaluable when creating DDS images of data. Interactively moving sliders to change the color, sizes, and animation parameters for each layer while seeing immediate results in the screen images, not only enables on to exert more precise control over the final image, but also reveals information in the data.

How do people experiment with the display parameters while creating visualizations of data with DDS? What are the patterns of adding multiple layers to the visualization image? Do people first create data mappings one layer at a time, waiting until all are created before looking at them overlaid in a single image, or do they add layers and readjust existing ones as they go? When do people add animation to the visualization?

What colors are used most often? In the first experiment I chose to test pastel colors because they seemed similar in that none dominated much in brightness or visual attraction. In creating visualizations of the SEM data sets I often choose jewel tones, and I seem to prefer dark blues and greens for some layers and bright yellows and pinks for other layers – when the parameter mappings are swapped I find the visualization unsatisfying in that I am unable to see the data as well as I think I should be able to. Some characteristics about the data seem to influence the color I prefer to use, but it is unclear what those characteristics are or if the color preferences are viewer-dependent.

Is color choice purely esthetic? Do esthetics choices improve or hamper data understanding, or is there no relationship between the two? Healey and Enns [2002] suggest that esthetics may in fact enhance understanding of the data. Is the final visualization satisfying because it is pretty or because it is easy to parse the layers visually?

Is there a difference between how a person uses the DDS visualization toolkit for data exploration and for data presentation? How long does it take to become proficient in creating images of data with DDS?

Figures 5.1 and 5.2 show different visualizations of the bat habitat range data created by two different people. Does a person create similar visualizations for different data sets, or are his or her spot size choices and color preferences data-dependent? People may find patterns that they prefer and always fall back on; I have a preference for saturated colors over pastels. Do different people create similar or distinctly different visualizations of the same data set?

### *Visual Perception Investigations with DDS Alpha-Blended Layers*

The experiments described in Chapter Three investigate a very small subset of the perceptual issues surrounding DDS. Although the results for the DDS alpha-blended layers were robust for up to nine data layers within an image, the results cannot be generalized to a wider class of data beyond simple geometric shapes and binary data.

Many characteristics of the underlying data influence how data is best displayed with DDS. The intricacy of boundary information and the presence of noise are just two examples. The example images presented in Chapter Two provide evidence that a wide range of data types can be effectively displayed with DDS, such as microscope, climate, and habitat range data. However, a detailed investigation would provide a better understanding of the limitations and strengths of DDS. Some questions are listed below.

How many discrete levels in the data can be conveyed through transparency? How does this change with multiple layers? How do hue and spot size affect the number of levels of transparency a person can see?

Human visual perception is hardwired to pick out boundary information in visual stimuli where boundaries may be difficult to see. Our perception enhances edges so we can accurately detect objects and navigate our world. Because boundary/edge detection is a specific mechanism of our visual perception system, it is interesting to ask how accurately people can see boundary information with DDS. How does this change with multiple layers? How much detail in the boundary can be seen? What if the boundary smoothly transitions, for example a gradient? What if the boundary is sharp? How do hue and spot size affect the perception of boundary details?

**Figure 5.1:** Bat habitat range for the following bats: California leaf-nosed (blue), California myotis (red), Eastern red (pink), Greater Bonneted (dark green), Indian myotis (dark blue), Mexican free-tailed (cyan), Pallid (light green), Rafinesque's big-eared (yellow, small spots), and Western red (yellow, large spots).

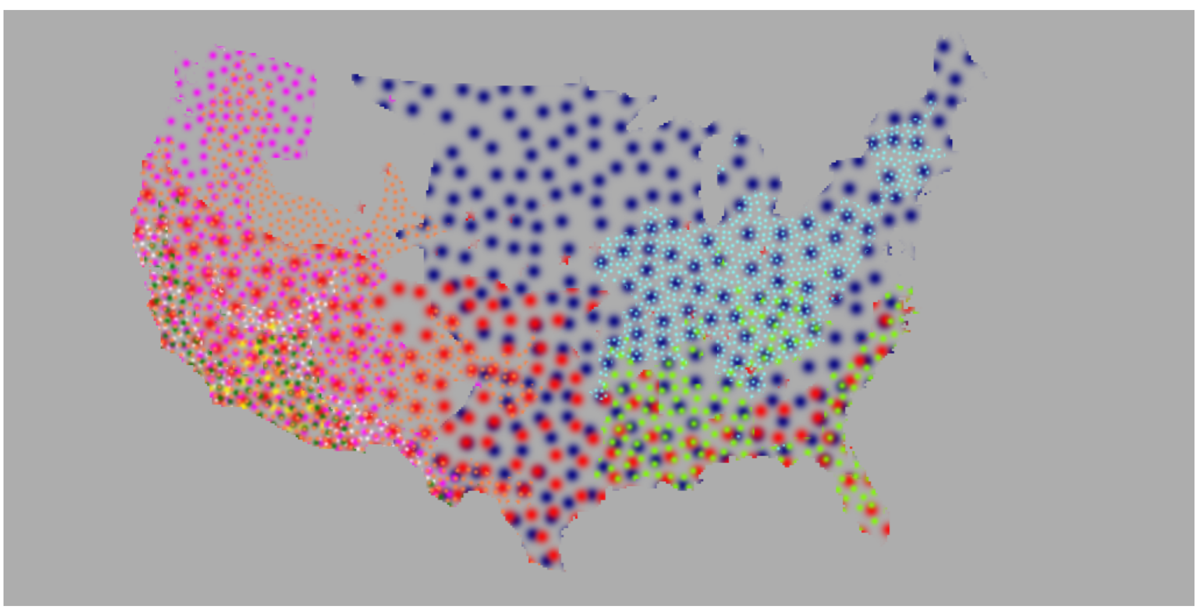

**Figure 5.2:** The same bat data, created by a different person. Data mappings are: California leaf-nosed (yellow), California myotis (purple), Eastern red (dark blue), Greater Bonneted (dark green), Indian myotis (cyan), Mexican free-tailed (red), Pallid (orange), Rafinesque's big-eared (light green), and Western red (white).

## Extending DDS to Higher Dimensional Data Visualization

### *Displaying DDS alpha-blended Layers on 3D Surfaces*

The experimental evaluation of DDS provides evidence that the DDS alpha-blended layers can be displayed on a non-planar surface with little surface interference with the DDS layers. We know this because the DDS bump-mapped layers did not interfere with the DDS alpha-blended layers. It is less clear whether the DDS alpha-blended layers would interfere with the perception of the underlying surface – task performance for the DDS bump map was affected by the distractor layers, but I believe that the problem was caused by applying multiple layers of bumps on top of other bumps and that there would not be a problem with one underlying surface.

There are several examples of data that would be better viewed on the native terrain instead of on a flat surface. The climate data set is a good example. In Chapter One, which presents frost frequencies over North and South America for the months of February and August averaged during the ten year period between 1981 and 1990, the images show frost even in the summer months in areas of high elevation, such as the Andes and Rocky Mountains ranges. The images would be more informative if the DDS layers were displayed directly on an elevation representation. Figures 5.3 and 5.4 show four climate variables – radiation (W/square meter) in yellow spots, vapor pressure (hPa) in red spots, wet-day frequency (days) in green spots, and precipitation (millimeters) in blue – all are displayed over a shaded relief topographic map. Figure 5.3 shows the averages for the month of February; Figure 5.4 shows the averages for the month of August.

The bat habitat data would also be more informative if displayed over terrain, because it would then be possible to see how the different species distributions relate to elevation. The AFM provides another good example, where the data is collected over a nanometer-resolution surface and would be better displayed directly on the original surface.

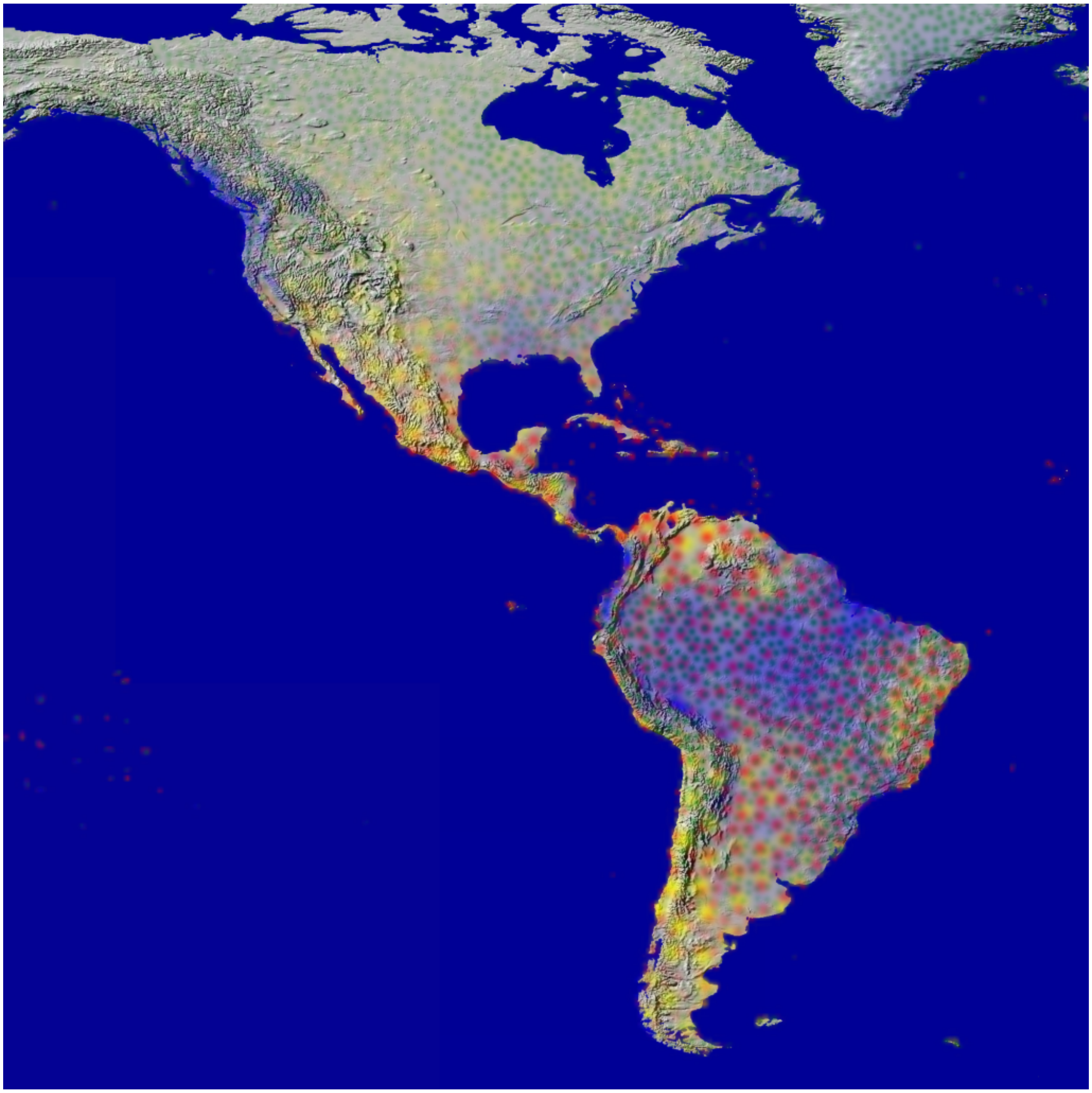

**Figure 5.3:** Climate data for February displayed over a shaded relief of a topographic map. Radiation (W/square meter) is shown in yellow spots, vapor pressure (hPa) is shown in red spots, wet-day frequency (days) is displayed with green spots, and precipitation (millimeters) is shown in blue.

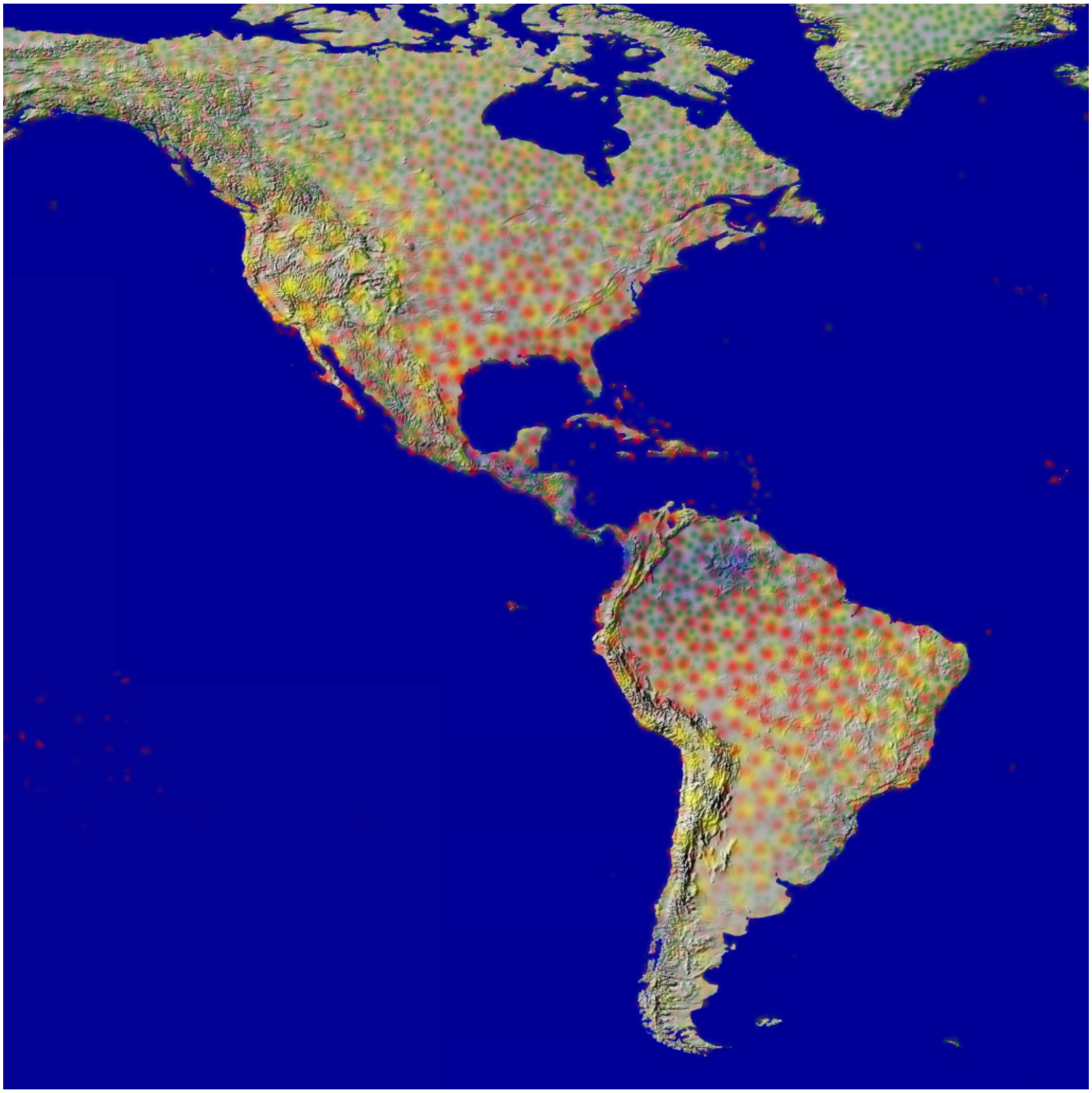

**Figure 5.4:** Climate data for August displayed over a shaded relief of a topographic map. Radiation (W/square meter) is shown in yellow spots, vapor pressure (hPa) is shown in red spots, wet-day frequency (days) is displayed with green spots, and precipitation (millimeters) is shown in blue.

***Visualization of Volume Data with DDS Alpha-Blended, Colored, 3D Gaussian Blobs***

Would the spatial sampling technique of DDS work for 3-dimensional volume data? Instead of 2D Gaussians sampling 2D spatial data, 3D Gaussian blobs could be used to sample 3D volume data. As the transparency of each pixel inside a Gaussian spot is determined by the value of the variable at that pixel, the transparency of each voxel inside a blob would be determined by the value at that voxel. Where the data is high, the blob would be mostly opaque, and where the data is low, it would be mostly transparent. Data outside the voxels sampled by the Gaussian blobs would not be displayed, allowing other variables to show through. This technique could be used to display multiple intersecting 3-dimensional fields. The blobs would have different colors to distinguish different variables measured on the volume.

Consider as an example a multi-valued atmospheric data set, such as would be produced by a computer weather simulation. At each point, one would have temperature, pressure, moisture content, wind direction, and wind velocity, etc. Each atmospheric variable would be sampled in 3D with a 3D array of Gaussian blobs. Within a 3D array of Gaussian blobs, each blob would have the same size and color. The spacing would depend on the size of the blobs within the array, as for the 2D array of Gaussians described in Chapter Two. The size, color, and sample spacing of the Gaussians would be different among the different 3D arrays of Gaussians to distinguish among variables. Sampling would be as for the 2D case: where the value of the volume variable was low the Gaussian blob would be transparent, where the value of the volume variable was high, the Gaussian would be opaque.

The number of different volumes we could see in such an image would likely be fewer than the number of layers for the 2D case. User interaction would be important when viewing the volume visualization – the viewer would have to be able to rotate the volume and view it from different angles. Animation of the underlying 3D Gaussian arrays through the volume would also be important to help the viewer see changes in the data values and to discriminate among variables.

## Augmenting the DDS Visualization Toolkit

### *Automatically Generating Best-Fit Mapping of DDS Alpha-Blended Display Parameters to Data Set*

There are two ways the DDS visualization toolkit could be augmented to help users with the task of finding the best mapping from data to display. The first method is to develop an expert system, similar to the work reported in [Healey, Amant, and Elhaddad, 1998], that would search through the space of visualizations, creating potential data-to-feature mappings and automatically evaluating the result based on perceptual guidelines to limit visual interference. The system they describe begins by asking the user to define the characteristics of their data and the type of analysis they would like to perform. Healey et al.'s work provides an excellent example of how to make the user's task easier as well as how to better teach the user about important perceptual cues relevant to visualization.

A similar method would be the development of a genetic algorithm-type search of the DDS data visualization parameter space to produce visualizations of data. Genetic algorithms are described in [Sims, 1991]; they provide a way of searching large feature spaces where the search is guided by fitness functions defined by the user. Genetic algorithms could be used to search through the space of spot colors, densities, and sizes as well as layer orders for a given data set and the user would then select images that either show the data well, are esthetically pleasing, or both. The process would repeat, using the selected images to guide future choices of data mappings.

Even simple assistance could be useful if it were automatic. For example, when creating a multilayer image I often want to swap the order the layers are applied, while keeping all other parameter settings the same. This is something an automatic visualization could easily do that would make the task easier. If I were presented with a series of images that differed only in the order the layers were applied I could easily select the one I preferred and proceed with creating the final image. This is just one example of how the parameter space can be explored with automatic help from an intelligent toolkit.

## Final Remarks

One question remains: Is the DDS visualization technique useful? I have shown that the technique works for binary data with simple shapes, and that it is significantly better than viewing the same data in separate side-by-side images. I have shown examples of DDS images on a variety of data sets, all of which have unique characteristics. What I would like to see next is people using DDS in practice, hearing what they think, and discovering new results when they put it to uses I never expected.

This dissertation contains my observations, thoughts, and beliefs about how to display multivariate scientific data. Many questions remain unanswered, and many new questions arise from the work presented here. The path of this research leads not to one answer, but to a fork in the road from which many trails begin.