On Histograms and Isosurface Statistics

Hamish Carr *Member, IEEE*, Brian Duffy and Barry Denby

Abstract— In this paper, we show that histograms represent spatial function distributions with a nearest neighbour interpolation. We confirm that this results in systematic underrepresentation of transitional features of the data, and provide new insight why this occurs. We further show that isosurface statistics, which use higher quality interpolation, give better representations of the function distribution. We also use our experimentally collected isosurface statistics to resolve some questions as to the formal complexity of isosurfaces.

Index Terms—histograms, isosurfaces, isosurface statistics

1 Introduction

Scientific and medical visualization commonly represents physical phenomena as continuous functions sampled over a spatial domain, then reconstructed by interpolation, either with a filter kernel or a geometric mesh. In either case, interpolation implicitly applies the spatial relationship between sample points to determine the function value at previously unsampled points. This interpolation method is then applied when visualizing the data.

Independent of the method chosen for visualizing spatial data, the fundamental task is often to identify important isovalues. The importance of an isovalue is usually measured using statistics of the sample points, most commonly the well-known *histogram*. This technique visualizes the distribution of the samples over the function's *range* by binning the samples and displaying a bar graph of the number of samples in each bin.

Histogram computation, however, assumes independent samples with no inherent relationships. In contrast, in spatial sampling of image or volumetric data, the close spatial relationships of the samples imply functional relationships, i.e. dependent samples. In Section 3, we show that the histogram is formally equivalent to the *nearest neighbour* interpolant, which is widely recognized as the worst possible interpolant that can be chosen. We believe that this relationship has gone unnoticed principally because of the ubiquity of the histogram as a statistical tool, and because the histogram was developed in contexts such as population statistics, where no meaningful spatial interpolant exists.

In Section 4, we show how to remedy the defects of the histogram by substituting isosurface statistics, which give substantially better representation at little additional cost. We demonstrate these defects, and the superior quality of isosurface statistics, by comparing histograms with isosurface statistics for a variety of real data sets made available on the internet.

In Section 7, we use our isosurface statistics to confirm some existing estimates of isosurface complexity. In particular, we show that Itoh & Koyamada's estimate [8] of $O(N^{2/3})$ is an underestimate, and that the observed relationship is closer to $O(N^{0.82})$. We also confirm a previous observation [4] that Marching Cubes generate an average of 2.05 triangles per cube.

2 Previous Work

Histograms are one of the oldest techniques known for displaying statistics [7]: the term was coined in 1892, but the technique may have been used earlier. Fundamentally, the histogram is a bar graph used to represent the distribution of function values in a population. In this bar graph, the independent variable represents the possible values of a set of observations, and the dependent variable represents the number of observations with a given value.

In computer graphics, image processing and visualization, histograms are often computed for the pixel (or voxel) values in a sampled data set. This information is then used for histogram equalization [6], manual transfer function construction [6, 9] and automated detection of significant isovalues [14, 15].

We show in Section 3 that histograms assume nearest neighbour interpolation: we observe that some authors compute histograms, while others have computed isosurface statistics based on specific interpolants. None, however, appears to have considered the relationship between the interpolant assumed for visualization, and the interpolant assumed for assessing function distributions.

Bajaj, Pascucci & Schikore [1] displayed isosurface statistics and topology in their *contour spectrum* to give users cues to interesting isovalues. These isosurface statistics were computed precisely in piecewise polynomial form based on an assumed linear barycentric interpolant over simplicial cells.

Carr, Snoeyink & van de Panne [3] extended these statistical computations to individual contours but used discrete approximations of their statistics by counting and summing individual sample values: this can be viewed as using nearest neighbour interpolation for their statistics, i.e. using histograms.

Pekar, Wiemker & Hempel [14] suggested using discrete isosurface statistics to supplement histograms for detecting significant isovalues. In addition, these authors showed how to compute the Laplacian-weighted histogram to find isovalues at which significant boundary effects occurred. Given the discrete nature of the Laplacian computation described, the effect of this is to convolve the image with a small interpolation kernel, then compute the histogram. Again, this is performed on individual voxels, and neither assumes nor employs an interpolant when generating statistics.

Tenginakai, Lee & Machiraju [15] used multi-dimensional histograms based on discrete computations of local higher-order moments, but used single isovalue bins without spatial interpolation.

Thus, although both isosurface statistics and histograms have been used to identify significant isovalues and other properties of data sets, individual authors have tended to use one or the other but not both, and the relationship between these two approaches has gone unexamined. In this paper, our principal aim is to examine this relationship, and to argue that isosurface statistics are a better representation of function distribution than histograms.

Isosurface statistics are also used to analyse isosurface extraction algorithms, which exploit the fact that k, the output cost of rendering, is significantly less than the input size N. Itoh & Koyamada [8] estimated that $k = O(N^{2/3})$ based on a dimensionality argument. In Section 7, we show that this is an underestimate, and that there appears to be some dependence on the type of data. We also confirm the estimate given by Carr, Theußl & Möller [4] for the average number

- *Hamish Carr is with University College Dublin, E-mail: hamish.carr@ucd.ie*
- *Brian Duffy and Barry Denby are also with University College Dublin*

Manuscript received 31 March 2006; accepted 1 August 2006; posted online 6 November 2006.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

of triangles rendered per active cube by Marching Cubes.

3 Histograms as Distributions

We begin by reviewing definitions of the histogram and the function distribution, then use these to show that the histogram is the function distribution for the nearest neighbour interpolant, but converges to the correct solution in the limit case.

3.1 Histograms

Histograms represent the distribution of a discrete set $\{h_i\}$ of samples, originally population data or other collections of isolated measurements, for which relationships between individual samples are generally not known. Formally, the histogram can be defined by:

$$H(h) = \sum_i \delta(h - h_i) \tag{1}$$

where δ is the *discrete* Dirac delta defined as:

$$\delta(x) = \begin{cases} 1 & x = 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

3.2 Continous Function Distributions

Image processing and visualisation assume a continuous function $f : \mathscr{D} \to \mathbb{R}$, where $\mathscr{D} \subset \mathbb{R}^d$. For such a function, we wish to compute the *distribution function*:

$$\pi_f(h) = \int_{\mathscr{D}} \delta(h - f(x))dx \tag{3}$$

where δ is the *continuous* Dirac delta defined as:

$$\delta(x) = \begin{cases} \infty & x = 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Employing the sifting property [2] of the Dirac delta, we can reduce Equation 3 to the following:

$$\begin{aligned} \pi_f(h) &= \int_{f^{-1}(h)} 1\, dx && (5) \\ &= Size(f^{-1}(h)) && (6) \end{aligned}$$

which computes the size of the *inverse image* $f^{-1}(h)$: i.e. the the set of points for which the Dirac delta is 1.

3.3 Point Sampling and Function Reconstruction

In practice, f is *sampled*, or evaluated, at a set of points $\{x_i\} \in \mathscr{D}$. The value $f(x_i)$ may be the function value at x_i or an integral over a small region, but is usually assumed to be the actual value at x_i, i.e. a point sample. The function distribution is then found by applying Equation 6 to the function reconstructed from $\mathscr{S}$ by convolving an interpolation kernel with the samples or by performing geometric interpolation over a polyhedral mesh covering the domain $\mathscr{D}$.

3.4 Histograms as Nearest Neighbour Interpolant

The simplest (and worst) form of interpolation used is the *box filter* or *nearest neighbour* interpolant, for which the reconstructed function F can be defined in terms of Voronoi cells [16]:

$$F(x) = \begin{cases} f(x_i) & x \in Vor(x_i) \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $Vor(x_i)$ refers to the Voronoi cell defined by the point x_i.

Although formally discontinuous, F is close enough that the distribution function can still be computed with Equation 6. Given Equation 7, we can see that for any h, the inverse image $F^{-1}(h)$ is the union of all Voronoi cells defined by samples with the value h:

$$F^{-1}(h) = \bigcup_{f(x_i)=h} Vor(x_i) \tag{8}$$

and using the Dirac delta to select only the Voronoi cells with the desired value, we can compute the size of the region as follows:

$$\begin{aligned} Size(F^{-1}(h)) &= \sum_{f(x_i)=h} Size(Vor(x_i)) && (9) \\ &= \sum_i \delta(h - f(x_i)) Size(Vor(x_i)) && (10) \end{aligned}$$

While true for any arbitrary set of samples, the most common sampling is a regular lattice, usually cubic or body-centred cubic [4]. For such lattices, the Voronoi cells are regular and, barring boundary samples, uniform in size. For a given size ζ, we then get:

$$\begin{aligned} \pi_F(h) &= Size(F^{-1}(h)) \\ &= \sum_i \delta(h - F(x_i)) Size(Vor(x_i)) \\ &= \sum_i \delta(h - F(x_i))\zeta && (11) \\ &= \zeta H(h) && (12) \end{aligned}$$

and the histogram is, up to a constant factor ζ, the distribution function for the nearest neighbour interpolant over regular lattices.

While the histogram is a poor representation for regular samples, it is an even worse representation for irregular samples data, since it treats all samples as having equal sized Voronoi cells. This over-emphasizes densely-sampled regions and under-emphasizes sparsely-sampled regions.

3.5 Convergence of Histograms

Recall that the integral of a function f over a region R can be defined by taking a set of patches P_ζ of size $\leq \zeta$ for any given $\zeta > 0$ such that the R is the union of the patches, and computing:

$$\int_R g(x)dx = \lim_{\zeta \to 0} \sum_{p \in P_\zeta} \zeta g(p) \tag{13}$$

i.e. a limit sum of boxes with bases p of size ζ or less and whose heights are the function values $g(p)$ at the centre of the patches.

For any $\zeta > 0$, let $\mathscr{S}_\zeta$ be a set of samples from f whose Voronoi cells are of size ζ. For regular cubic lattices in three dimensions, this requires a sampling distance of $\sqrt[3]{\zeta}$. Since each Voronoi cell has size $\leq \zeta$ and their union is the entire domain $\mathscr{D}$, we can use them as the patches $\mathscr{P}_\zeta$ required in Equation 13.

Substituting Equations 13 and 10 into Equation 6, we get:

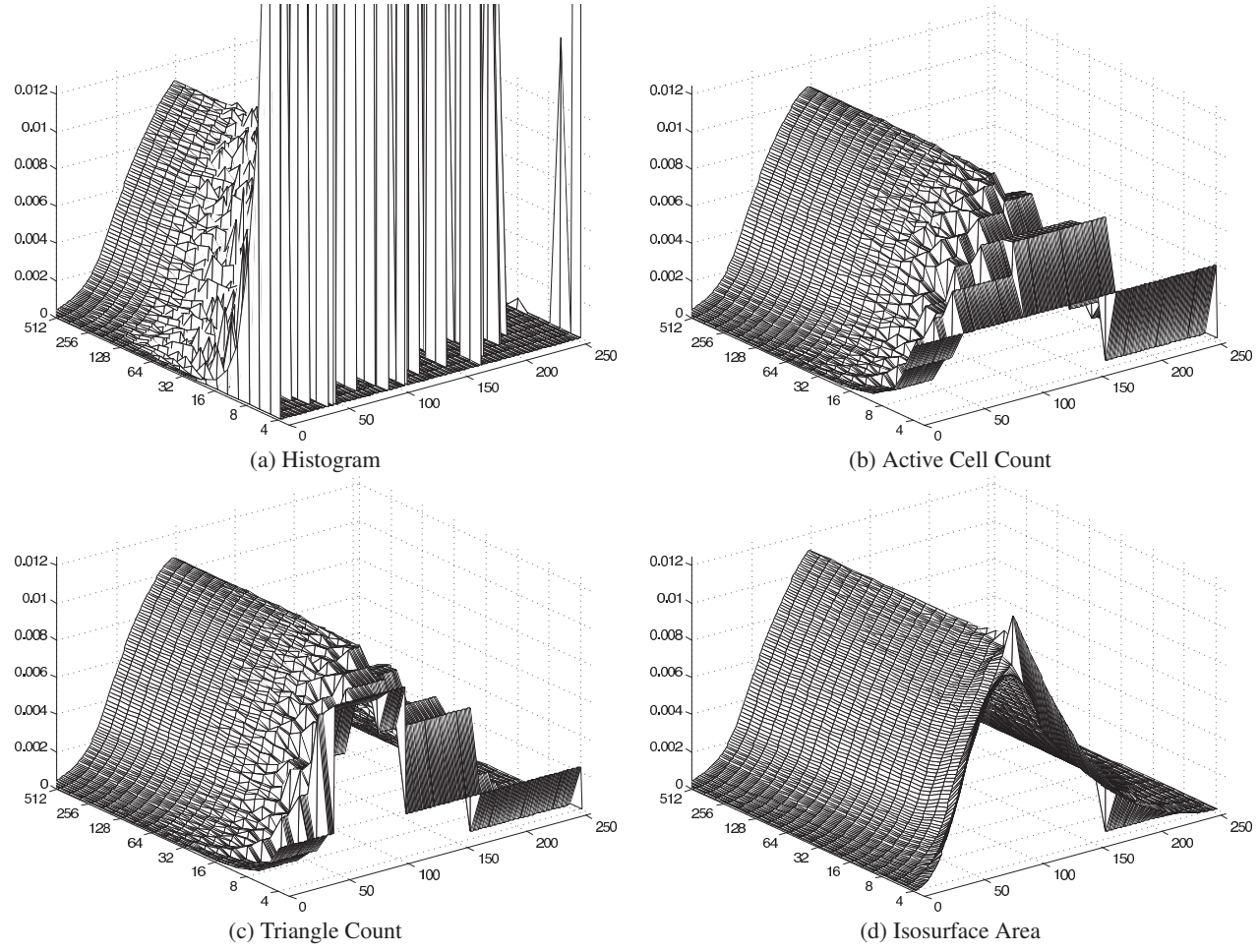
$$\begin{aligned} \pi_f(h) &= \int_{\mathscr{D}} \delta(h - f(x))dx \\ &= \lim_{\zeta \to 0} \sum_{\mathscr{P}_\zeta} \delta(h - f(x))\zeta \\ &= \lim_{\zeta \to 0} \pi_{F_\zeta}(h) \\ &= \lim_{\zeta \to 0} \zeta H_\zeta(h) && (14) \end{aligned}$$

proving that the histogram converges to the function distribution for sufficiently small sampling distances.

We have now seen that the histogram is formally equivalent to the function distribution for the nearest neighbour interpolant, but that it converges to the correct solution in the limit. In the next section, we see how *fast* it converges - i.e. how good it is in practice.

4 Isosurface Statistics as Distributions

In the previous section we showed that the histogram represents the function distribution imperfectly because it is based on the nearest neighbour interpolant. Ideally, we would like to compute the correct distribution function for any interpolation function chosen. We will show that this is equivalent to the task of computing the areas of



(a) Histogram (b) Active Cell Count

(c) Triangle Count (d) Isosurface Area

Fig. 1. Relative Quality of Spherical Function Distribution with Increasing Sample Resolution. All plots converge to the correct solution as the number of samples increases. Histograms give the poorest representation, while area gives the best representation, with triangle and cell counts nearly as good for all but the coarsest representations.

all possible isosurfaces, which Bajaj, Pascucci & Schikore [1] have demonstrated to be tractable for simplicial meshes.

Recall from Equation 6 that the distribution function for a value h is the size of the inverse image $f^{-1}(h)$. Conveniently, this inverse image is also the formal definition of an *isosurface* in three dimensions. Ideally, the reconstructed function is a Morse function - i.e. one with no *flat* regions, and $f^{-1}(h)$ is a $(d-1)$-manifold embedded in the d-dimensional domain $\mathscr{D}$. For $d=3$, the isosurface is a $2-$ manifold, whose surface area is the required size. Although the ideal case is rare, we can approximate it with symbolic perturbation of the sample values [5] or of the isosurface itself [10].

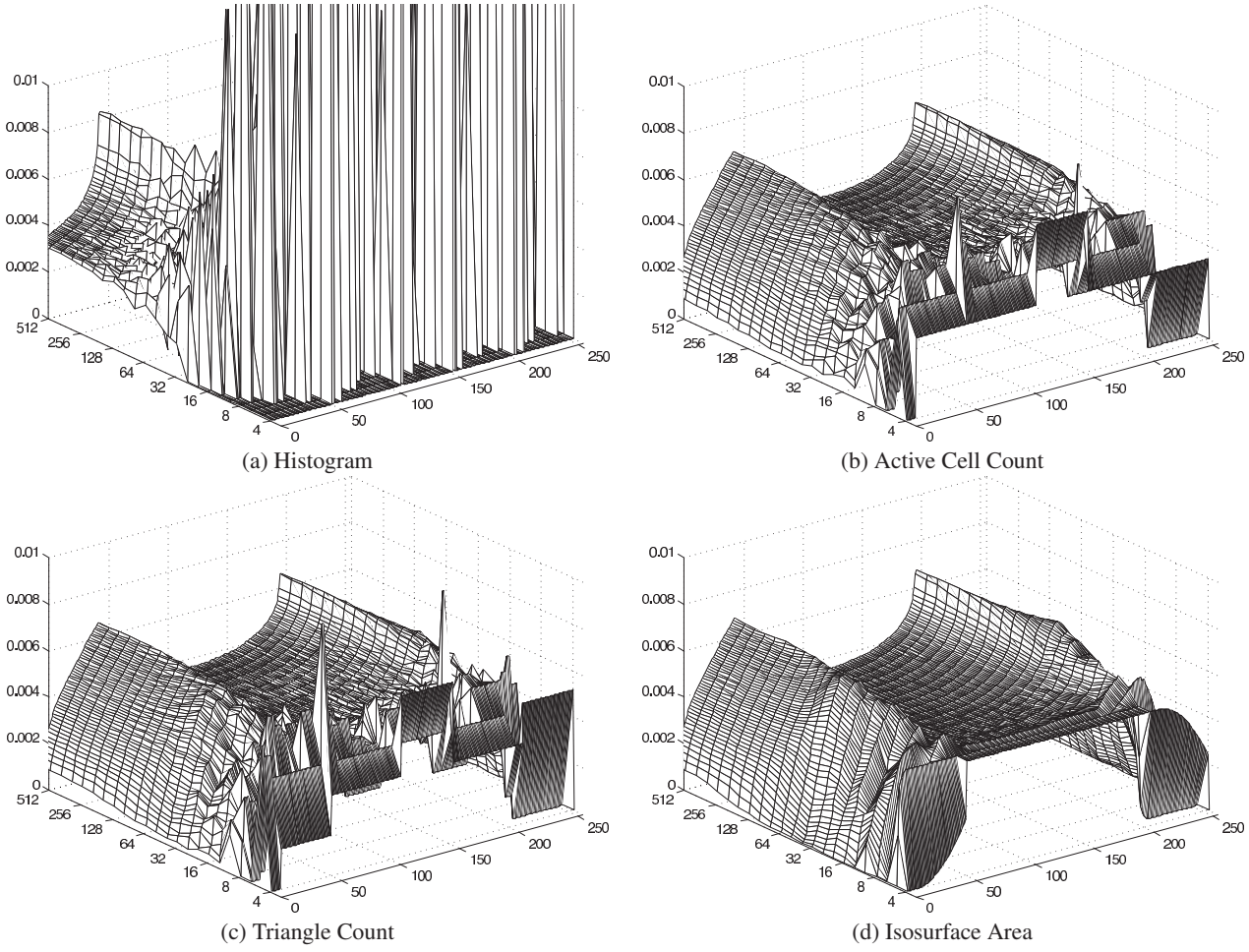
Thus, simply by computing the surface area of the isosurface at each isovalue, we can get the correct distribution function. Bajaj, Pascucci & Schikore [1] have shown that for simplicial meshes, the surface area is a piecewise quadratic polynomial in the isovalue h, whose coefficients are the sums of individual coefficients for each cell in the mesh. Relying on this property (decomposability), they also showed that the polynomial pieces can be computed by sweeping a hyperplane through the function from high to low isovalues, updating the coefficients as each sample point is swept past.

This approach is not easy to extend to more complex interpolation functions. For example, the bilinear interpolant in 2-D gives hyperbolic contours [13], whose size can only be computed with logarithmic terms [3]: trilinear and higher order interpolants are worse yet. Instead, we compute the area of the triangulated isosurface approximations generated by Marching Cubes [10] for each isovalue. For simplicity, we computed these directly, but where performance is an issue, polynomials can be computed for each cell by subdividing the regions bounded by the isosurface into tetrahedra, then using the method given by Bajaj, Pascucci & Schikore [1].

For convenience, we used the Marching Cubes cases of Montani, Scateni & Scopigno [12], which generate watertight triangulated approximations of the correct isosurface. Although not as accurate an approximation as the trilinear solutions, our results are considerably better than histograms. Moreover, for regular datasets, the triangles generated are fairly uniform in size, so the area computed can also be approximated by counting the number of triangles in the isosurface, or even the number of cells intersected by the isosurface. Although we collected these statistics principally for the purposes of Section 7, we observed that they were superior to histograms as a distribution and nearly as cheap to compute.

5 Comparison of Analytic Distributions

Once we have identified that isosurface statistics are theoretically better than histograms, the next step is to compare them for analytically defined functions by sampling at progressively higher resolutions and comparing their histograms. We chose a spherical distribution $f(x,y,z) = 255 - \frac{255}{\sqrt{3}}\sqrt{x^2+y^2+z^2}$ (Figure 1), and the Marschner-Lobb function (Figure 2). For these datasets, we took $2^{n/3}$ samples in each dimension, resulting in datasets with sizes ranging from $3\times3\times3$ to $512\times512\times512$ such that each dataset had twice as many samples as the previous one. For each resolution, we then computed the following statistics:



(a) Histogram (b) Active Cell Count

(c) Triangle Count (d) Isosurface Area

Fig. 2. Relative Quality of Marschner-Lobb Function Distribution with Increasing Sample Resolution. All plots converge to the correct solution as the number of samples increases. Histograms give the poorest representation, while area gives the best representation, with triangle and cell counts nearly as good for all but the coarsest representations.

1. The histogram
2. The number of cells intersected by each isosurface
3. The number of triangles generated for each isosurface
4. The total area of the triangles generated

Since each of these represents a distribution, we normalize each statistic for a given resolution by dividing the value through by the sum of the values in all bins.

The isosurfaces of the spherical distribution are concentric spheres, so we expect the function distribution to increase steadily as the isovalue decreases until the isosurfaces touch the boundary of the sampled domain. Thereafter, the isosurfaces are incomplete spheres, and we expect the distribution to decrease again. Since this distribution is a simple function with low-frequency information, we would hope to get good results even at coarse resolutions.

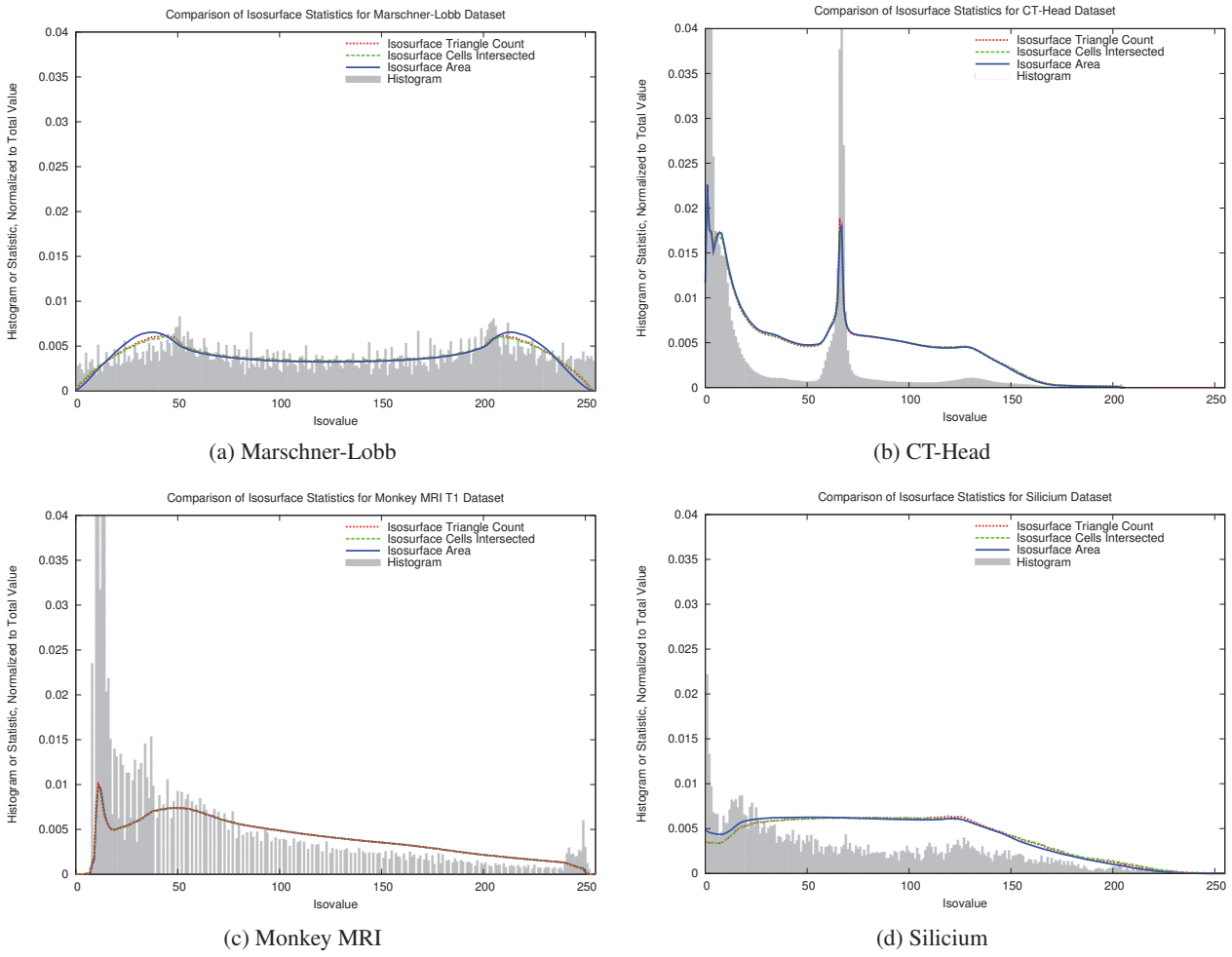
As we see in Figure 1, the histogram is a poor representation of this distribution at the coarsest resolutions, but improves with increased resolution until it approximates the correct distribution quite well at resolutions of 64^3 or higher. In comparison, however, the isosurface area computation is a good approximation even at resolutions as low as 4^3, while triangle counts and cell intersection counts are satisfactory from roughly 16^3.

This spherical dataset, although a useful illustration, is a very simple function with principally low-frequency information. We therefore repeated the experiment with the more complicated function introduced by Marschner & Lobb [11]. In Figure 2, we see the same trends: the isosurface area is the most accurate representation, with the histogram being the worst. Moreover, we can see that at the resolution of 40^3 chosen by Marschner & Lobb to test frequencies near the Nyquist limit, the histogram has not yet converged sufficiently to give a good representation of the signal.

6 Comparison of Experimental Distributions

In Section 5, we showed that histograms are poor representations of analytic function distribution compared to isosurface statistics such as area. To determine how well histograms represent typical datasets, we tested 94 volumetric datasets downloaded from the internet, including simulation data, medical data and other experimentally acquired data. We quantized each dataset to 8 bits for consistency, then computed the same statistical measures as in the previous section. Again, we normalized each statistic to its sum over all bins, and truncated all graphs at a common value of 0.04. While it is not possible to show all 94 graphs generated, some themes were apparent, and we show four typical graphs in Figure 3.

We start with the Marschner-Lobb dataset in Figure 3(a), since we showed in the last section how poorly the histogram approximates the distribution. In Figure 2, we saw that the function distribution in the limit has two peaks separated by a wide valley. As we see in Figure 3(a), the histogram at 40^3 is only a crude approximation, but isosurface area (the blue dotted line) gives a nice smooth representation of the distribution we expect, with two clearly defined peaks. Moreover, we also notice that despite some raggedness, both the isosurface cell intersection count and the triangle count give quite good approx-



(a) Marschner-Lobb (b) CT-Head

(c) Monkey MRI (d) Silicium

Fig. 3. Comparison of Histograms and Isosurface Statistics. For the Marschner-Lobb dataset, we know the correct (limit) distribution from Figure 2, and we observe that the isosurface statistics are a much closer approximation than the histogram. We also notice, however that while the computed area is clearly the best representation, both cell intersection count and triangle count are nearly as good. Consistent problems of the histogram include under-representation of transitional values, omission of side peaks, and introduction of erroneous peaks.

imations of the curve. We believe that this raggedness is because the Marschner-Lobb function is highly symmetric and regular. This implies that every time one triangle is added, 3 other symmetric triangles are added, magnifying errors in the representation of the distribution.

Figure 3(b) compares the statistics for a CT scan of the head (CT-Head dataset). In this, the histogram is distorted by a huge peak at 0 (the region outside the head), which makes all other features disappear. Even once this has been eliminated by truncation, however, the secondary peak at around isovalue 10 is completely missing from the histogram, and the small peak at around 130 is barely visible. Moreover, the transitional regions between peaks are under-represented in the histogram. All of these problems were common to most of the datasets, and in particular to experimentally acquired datasets such as medical CT and MRI scans, but were also visible in simulated data. Finally, we notice that the isosurface triangle count and cell intersection count are in fact quite close approximations of the area, presumably because noise in the data smoothes out the effect noted in Marschner-Lobb.

Figure 3(c) shows statistics for a T1 MRI scan of a monkey. In this dataset, the secondary peak visible at isovalue 50 is not visible in the histogram, and the histogram shows what appears to be a peak above 235, while the isosurface area shows that the function is actually decreasing smoothly over these isovalues.

Figure 3(d) shows the comparison for a simulated dataset of a silicon crystal (Silicium). As we see here, the histogram is dominated by a large peak of low values which have very limited effect on the overall distribution of the function, while the isosurface statistics correctly represent the importance of medium isovalues. And, although a secondary peak is visible in both the histogram and the statistics at around 120, it is more prominent in the histogram and shifted slightly to the left. Although there are differences in detail, all three of the isosurface statistics are relatively smooth, although there is some raggedness in the isosurface cell intersection and triangle counts.

6.1 Trends

In general, we observed the following trends in our graphs:

1. Histograms systematically underestimated transitional regions, especially in experimental data sets, but also in simulated data sets.

2. Histograms commonly missed secondary peaks in the function distribution, and sometimes even displayed spurious peaks.

3. Histograms of simulated data were frequently ragged, as it was a largely a matter of chance whether the samples clustered on one isovalue or were spread among many. This effect was uncommon for experimental datasets, in which the presence of measured noise tended to smooth out the histogram - see Figure 3(b) for an example.

4. Although the area measurement was clearly smoothest in all datasets, the isosurface triangle count and cell intersection count,

which are much cheaper to compute, were generally nearly identical to it.

We conclude that, in any application where the distribution of spatially interpolated functions is useful information, the histogram should be replaced by isosurface statistics, and that even a relatively inaccurate measurement such as the number of cells intersected by each isosurface is a far better representation of the distribution.

7 Statistical Complexity of Isosurfaces

Since we were collecting isosurface statistics in order to examine the quality of histogram representations of function distributions, we also took the opportunity to consider some questions relating to the formal complexity of isosurface extraction. In particular, we wished to test the estimate by Itoh & Koyamada [8] of isosurface complexity and the estimate of Carr, Theußl and Möller [4] of the average number of triangles extracted per Marching Cube.

7.1 Isosurface Complexity

An important issue in isosurface extraction is the formal complexity of algorithmic improvements, which generally rely on output-sensitive analysis, in which the rendering cost k, the number of triangles generated, is balanced against the pre-processing costs based on N, the number of input samples. Although pathological datasets such as checkerboards may have $k = \Theta(N)$, few researchers are interested in visualizing these datasets. Instead, the common experience is that $k << N$. To date, only one estimate of the relationship has been made: that of Itoh & Koyamada [8], who estimated that $k = O(N^{2/3})$ in the course of a paper on an algorithmic improvement. The sole justification advanced for this estimate was one of dimensionality: that as a 2-manifold extracted from a 3-manifold, the isosurface could reasonably be expected to have $O(N^{2/3})$ triangles. Other arguments can be advanced for other asymptotic relationships, but since the question is one of the actual isosurface sizes not theoretical models, we decided to collect data experimentally and test the hypothesis.

As described above, we collected 94 different datasets and computed isosurface statistics for each. For each dataset, we computed the average number of triangles (k) across all isovalues and compared it with the number of samples (N). If the correct relationship between these is the $O(N^{2/3})$ hypothesized, we should able to see this in a log-log scatterplot, in which the observations would cluster around a line with a slope of $2/3$. Figure 4(a) shows just such a plot, in which the input size N is shown on the x-axis and the output size k on the y-axis. When a least-squares line is fitted to the observations, the slope computed is approximately 0.82, which indicates that this dimensionality argument underestimates the typical result.

On closer inspection, we discovered that different types of data appear to have different relationships, with simulated data typically having a smaller exponent and noisy data having a higher exponent, as much as 1.05 in the case of medical data, as seen in Figures 4(b) to 4(d). While we find this observation interesting, we are aware that 94 datasets is not a large collection, and our classification into different types of data meant that we were producing scatterplots for 30 or fewer data points. Still, this is an indication that there may be subtle relationships that depend on the type of data, and which could be exploited for automated classification of data sets.

7.2 Average Triangle Counts

Carr, Theußl & Möller [4] estimated that Marching Cubes should generate 2.85 triangles per active cube, based on a probabilistic argument, but reported figures of roughly 2.0 triangles per active cube, based on a small sample of data sets. Since we had computed isosurface sizes and cells intersected, we took the opportunity to double-check this result across all the data sets we tested. For each dataset, we computed the average number of triangles per active cube across all isovalues, then averaged these results between all datasets. We obtained a result of 2.04312 ± 0.062791, confirming the previous result.

8 Conclusions

Histograms are often used to represent function distributions, but we have shown that they are poor representations for spatially interpolated data since the histogram gives the distribution of the nearest neighbour interpolant. We have shown that the correct method of representing function distributions is by computing the area of the isosurface under the chosen interpolant, but that approximations of this area, even crude approximations such as triangle counts and cell intersection counts, are significantly superior to histograms.

We have also tested the estimate of Itoh & Koyamada that isosurface complexity is $O(N^{2/3})$ in the size of the input, and reached the conclusion that it underestimates the experimental results of $O(N^{0.82})$. Finally, we confirmed the estimate of Carr, Theußl and Möller that the average number of triangles per active cube in Marching Cubes is 2.04 ± 0.06.

9 Future Work

We would like to expand our study to a larger number of datasets, but this is difficult without access to many sources of data, as we would not wish our results to be biased by a single source. We would also like to compute isosurface statistics more accurately for trilinear and higher interpolants, but recognize that this may be a complex task in its own right.

Finally, we would like to investigate the possibilities raised in our experiments on complexity, to see if different types of data have identifiably different statistical characteristics permitting automatic or semi-automatic analysis of arbitrary data sets.

10 Acknowledgements

Acknowledgements are due to those responsible for making volumetric datasets available over the Internet, including www.volvis.org. Acknowledgements are also due to University College Dublin and Science Foundation Ireland for supporting this research in the form of undergraduate research grants and research support grants. The authors would also like to thank the anonymous reviewers for their constructive suggestions.

References

[1] Chandrajit L. Bajaj, Valerio Pascucci, and Daniel R. Schikore. The Contour Spectrum. In *Proceedings of Visualization 1997*, pages 167–173, 1997.

[2] Ronald Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, 3rd edition, 1999.

[3] Hamish Carr, Jack Snoeyink, and Michiel van de Panne. Flexible isosurfaces: Simplifying and displaying scalar topology using the contour tree. Submitted for journal publication, 2006.

[4] Hamish Carr, Thomas Theußl, and Torsten Möller. Isosurfaces on Optimal Regular Samples. In *Proceedings of Eurographics Visualization Symposium 2003*, pages 39–48, 284, 2003.

[5] Herbert Edelsbrunner and Ernst P. Mücke. Simulation of Simplicity: A technique to cope with degenerate cases in geometric algorithms. *ACM Transactions on Graphics*, 9(1):66–104, 1990.

[6] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (2d ed.)*. Prentice-Hall Inc., Englewood-Cliffs, NJ, 2002.

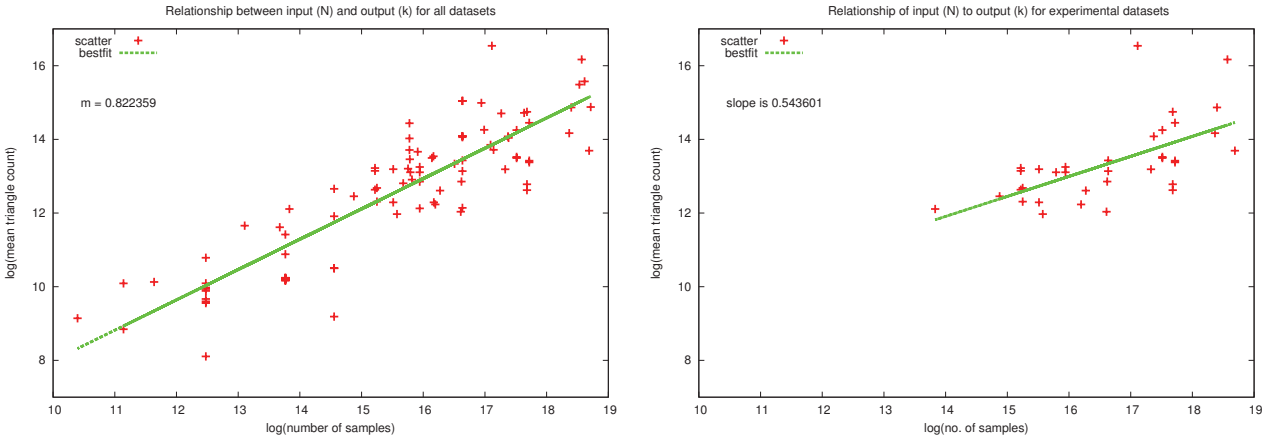
[7] Yannis Ioannidis. The history of histograms. In *Proceedings of Very Large Databases (VLDB) 2003*, 2003.

[8] Takayuki Itoh and Koji Koyamada. Isosurface Extraction By Using Extrema Graphs. *IEEE Transactions on Visualization and Computer Graphics*, 1:77–83, 1994.

[9] Joe Kniss, Gordon Kindlmann, and Charles D. Hansen. Interactive Volume Rendering Using Multi-Dimensional Transfer Functions and Direct Manipulation Widgets. In *Proceedings of Visualization 2001*, pages 255–262, 562, 2001.

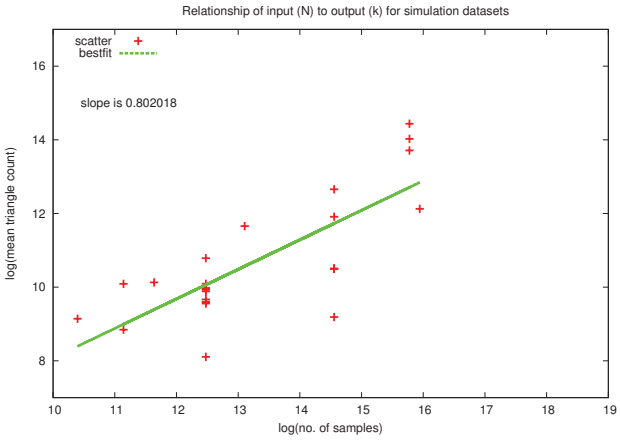
[10] William E. Lorenson and Harvey E. Cline. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. *Computer Graphics*, 21(4):163–169, 1987.

[11] Stephen R. Marschner and Richard J. Lobb. An Evaluation of Reconstruction Filters for Volume Rendering. In *Proceedings of Visualization 1994*, pages 100–107, 1994.

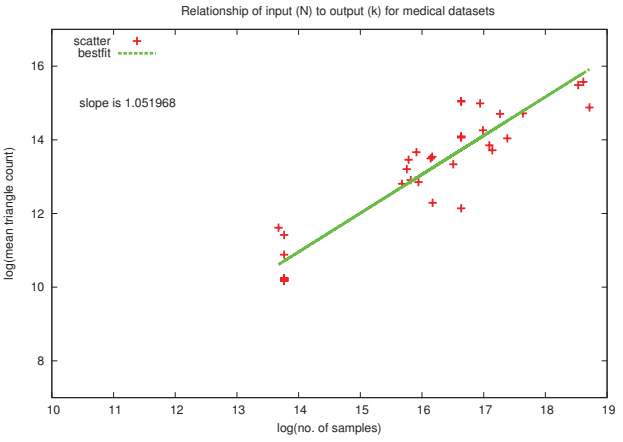


(a) All Data Sets

(b) Experimental Only



(c) Simulation Only



(d) Medical Only

Fig. 4. Plots of Input Complexity (N) versus Output Complexity (k). Despite the small number of observations, there appears to be a polynomial relationship between N and k.

[12] Claudio Montani, Riccardo Scateni, and Roberto Scopigno. A modified look-up table for implicit disambiguation of Marching Cubes. *Visual Computer*, 10:353–355, 1994.

[13] Gregory M. Nielson and Bernd Hamann. The Asymptotic Decider: Resolving the Ambiguity in Marching Cubes. In *Proceedings of Visualization 1991*, pages 83–91. IEEE, 1991.

[14] Vladimir Pekar, Rafael Wiemker, and Daniel Hempel. Fast Detection of Meaningful Isosurfaces for Volume Data Visualization. In *Proceedings of Visualization 2001*, pages 223–230, 2001.

[15] Shivaraj Tenginakai, Jinho Lee, and Raghu Machiraju. Salient Iso-Surface Detection with Model-Independent Statistical Signatures. In *Proceedings of Visualization 2001*, pages 231–238, 2001.

[16] G.M. Voronoï. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. premier mémoire: Sur quelques propriétés des formes quadratiques positives parfaites. *Journal für die Reine und Angewandte Mathematik*, 133:97–178, 1907.