5

Complex measurements made easy: morphometric analysis of artefacts using Expert Vision Systems

Irwin Rovner

(North Carolina State University, Raleigh, North Carolina, USA)

5.1. Introduction

In conventional archaeological analysis descriptions of object morphology normally combine a few standard size measurements with some qualitative or anecdotal shape terms. Tape measures and calipers limit us to measurements of length, width and thickness. Shape is usually described by picture-stereotypes — "hat-shaped, cigar-shaped, barrel-chested" — or by imprecise qualitative terms, such as, roundish, irregular, nearly smooth to slightly rough, indented to sinuous, *etc.* Often some combination of picture-stereotypes and qualitative terms, such as "reddish, broad and bulbous like my Uncle Harry's nose" is used. When built into classification and typology, such description panders to our penchant for thinking in simplistic stereotypes while conscientiously ignoring individual variation. We revel in arguments over the reality of "types" and in the assignment of individual pieces to type categories based on the universal principle that no two classification systems shall ever sort an assemblage of artefacts in the same way. If ever in doubt, employ the universally applicable category "other...", or create the type category with the oxymoronic title "atypical..."

Stereotypical procedures continue the obsolete logic of the era when biological organisms conformed to John Ray's divinely crafted immutable species or correlated to George Cuvier's perfect designs. Contemporary semantic substitutions, such as wild type, central tendency, and prototype improve nothing. The theoretical importance of individual variation, made compellingly clear by Josiah Wedgwood's grandson, Charles Darwin, has yet to dominate practical procedures of systematic classification in biology or archaeology. This is due partly to the ease of creating stereotypes, reinforced by the virtual inability of the few conventional measurements to quantify individual size and especially individual shape easily, accurately or precisely. However, with the advent of compact image processing computer hardware and robust stereological software, the era of limited manual measurement and conventional typology should come to an end quickly.

5.2. The ImagePlus II system

The case studies presented here were conducted using the (now-obsolete) ImagePlus II software package, designed by John C. Russ of North Carolina State University, which operates in an enhanced past-generation Apple II series personal computer. The current generation of this system, entitled Prism, operates exclusively on Macintosh personal computers and is available in the USA through Dapple Systems, Inc. Similar IBM-based systems are available from several companies such as Cambridge and Jandel.

The ImagePlus system has many functions and modes of image analysis applicable to aggregates, conglomerates, mosaics and maps. The most immediately amenable use of the system is in analysis of assemblages of discrete items, such as artefacts, subjected to morphological measurement of size and shape as a basis for classification and object recognition. To do this, ImagePlus acquires and digitises images generated by CCD TV camera, flatbed scanner, videotape, or video output from electron, laser or acoustic microscopes. The analogue images are displayed in a 512 x 512 pixel (picture element) TV monitor in either 256 levels of B&W grey scale or 16.3 million levels of pseudocolour. The program provides several menu-driven methods of image processing, editing and enhancement.

Grey scale (or pseudo-colour) discrimination is used to isolate objects of interest in an image field. Selected items (or areas) are converted from grey scale to binary (full black and full white) images. These may be subjected to further image processing and enhancement. The system offers a menu of 40 measured parameters up to 20 of which may be selected at one time for simultaneous measurement of a maximum of 254 discrete objects in a single field. The system also allow the operator to write additional, custom-designed measurement algorithms for the program. In reality, an image field usually consists of some 5 to 50 objects on which 10 to 15 different measurements will be taken. On average, a single skilled operator can generate from 2,000 to 5,000 measurements per hour. Data may be subjected to statistical testing at any time within the program with results graphically displayed.

5.3. Morphometry

Object morphometry and object recognition are achieved using five parameter categories: dimension of size, dimensionless ratio of shape, harmonic analysis, fractal dimension and topology (Russ, 1992, pp. 175–318).

ImagePlus offers both conventional and unconventional size measurements which often have subtle but critical differences when compared to manual measurements of the same name. "Length" in ImagePlus, for example, is the longest chord within the object perimeter regardless of object orientation (Fig. 5.1). This inelastic measurement is not dependent on arbitrary orientation on an X-Y co-ordinate grid which can result in systematic bias and error. Actually, ImagePlus also measures orientation-dependent values, X- and Y-Feret's (horizontal and vertical distances), as well. If the longest chord within an object's perimeter is

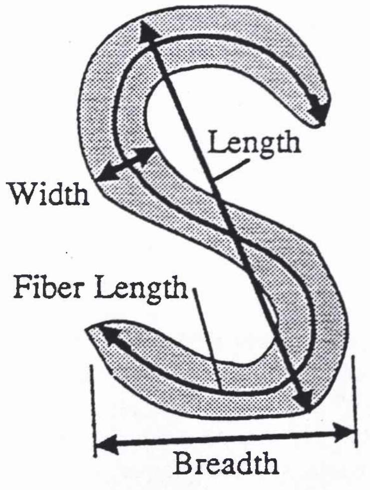


Figure 5.1: Examples of basic computer measurements of distance sizes.

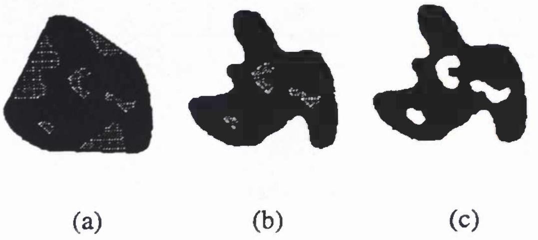


Figure 5.2: Three ways to measure area of an irregular object. (a.) The convex or "taut string" area; (b.) the externally occupied area with interior holes "filled"; and, (c.) the "true" area.

oriented along the X axis, then length and X-Feret are equal. Such conventions in manual measurement would be no problem if everything in archaeology were square or round or rectangular, but clearly artefacts and fossils are notoriously irregular in shape. This gives great significance to arbitrary orientation in manual measurement while the computer always measures the longest length dimension of an object regardless of irregularity or orientation.

The literal rules of computer measurement treat regular and irregular objects with equal facility. The computer can calculate the area of a digitised map of Great Britain as easily as it can that of a simple square. In addition the computer measures attributes that are difficult or impossible to derive manually, especially in cases of irregular objects. These include area, perimeter, "fibre length" (the total lineal distance down the centre line of an object), curl, radius of curvature and a host of other such dimensions (Fig. 5.1). When measuring the area of a doughnut, is the area of the hole included in the whole? If you wish to know how many doughnuts will fit on a plate, the answer is yes (Figs. 5.2a–b). If you wish to know how much sugar you need to coat the doughnut, the answer is no (Figure 5.2c). IMAGEPLUS can calculate it either way. The situation is no more complex for the computer if you are measuring Swiss cheese, the effect of yeast in a slice of bread, or the extent of osteoporosis in a bone specimen.

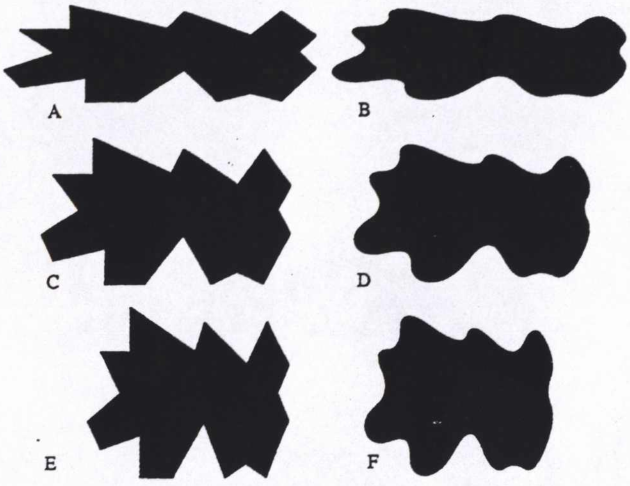
Dimensions of size are relatively easy to understand; but, for purposes of object recognition, they suffer from the fact that objects with obviously different shapes can yield the same size dimensions. To rectify this problem, IMAGEPLUS uses its several measurements of size to create dimensionless ratios that measure, that is, quantify, specific attributes of shape. This is possible because change in an aspect of shape directly influences changes in (ratios of) size dimensions as dependent variables. Archaeologists should be familiar with a shape factor of long standing; namely cephalic index or head shape. This is the dimensionless ratio of skull width divided by skull length $\times 100$. IMAGEPLUS arbitrarily calculates the reciprocal of this ratio under the name, 'aspect ratio', providing a simple and straightforward index of elongation.

The utility of shape factors in describing object attributes is seen in comparing values for 6 different objects of identical area in Fig. 5.3 which shows values of form-factor ($4\pi \times$ area / perimeter squared) and roundness ($4 \times$ area $/\pi \times$ length squared). For a perfect circle, the values of both shape factors equal 1. In squaring the perimeter measurement, form-factor values are very sensitive to changes in the shape of the perimeter but insensitive to elongation. Form-factor groups the three objects in the left column from the three in the right column. Roundness, which squares the dimension of length, is sensitive to elongation but insensitive to perimeter. Thus, roundness values group the horizontal adjacent pairs, while differentiating between the pairs vertically.

The shape factor known as "solidity" depends on the inverse relation of area to perimeter as object shape departs from a regular polygon or a circle, oval, *etc.* To calculate this, the computer fits a 36-sided polygon to each object, essentially creating a taut string around the object (Fig. 5.2a). The taut string is known as the "convex perimeter" which surrounds the "convex area". This permits calculation of ratios of true perimeter to convex perimeter to yield solidity as well as true area to convex area, another shape factor known as Convexity. If an object is a perfect polygon, the values of solidity and convexity equal 1. However, if an object with constant convex perimeter value becomes irregular, that is, has lobes and indentations, then true perimeter increases in direct proportion to irregularity while true area decreases. Ratio values departing from 1 indicate a quantitative measure of such irregularity.

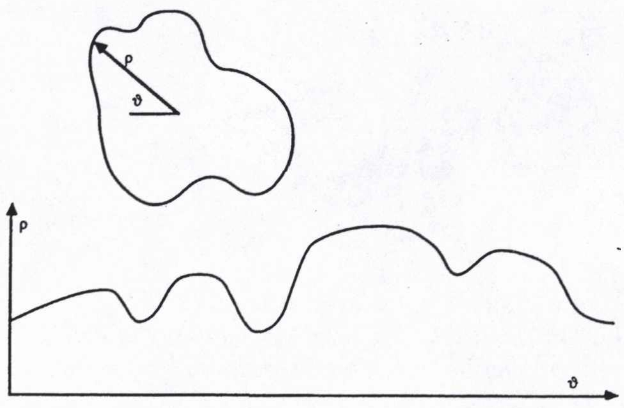
Dimensionless ratios effectively describe shape but at the level of discrete attributes. Object recognition requires selection and simultaneous use of multiple shape factors, often in conjunction with size dimensions, to create a classification system. However, with the easy availability of multi-variate statistics and discriminant function analysis, this is a trivial problem.

Harmonic analysis (Fig. 5.4) uses trigonometric functions and vectors to unroll a shape into a single mathematical expression of that shape which is then subjected, for example, to Fourier analysis. This is an enormously effective and successful way to quantify a shape, but it is limited by the fact the people do not easily "see" what the expression means. Since the thinking of people does not relate well to such coding, comparisons of expressions are not



Formfactor		*Roundness*	
A 0.311	B 0.447	A 0.260	B 0.278
C 0.332	D 0.467	C 0.497	D 0.502
E 0.318	F 0.477	E 0.623	F 0.649

Figure 5.3: Six objects of equal area grouped and separated by shape attributes using form-factor and roundness values.



$$\rho(\vartheta) = a_0 + a_1 \cos(\vartheta) + b_1 \sin(\vartheta) + a_2 \sin(2\vartheta) + b_2 \cos(2\vartheta) + \ldots$$

Figure 5.4: Harmonic analysis: shape plot of an "unrolled" object.

easily made. In fact, effective analysis is wholly mathematical and computer intensive often beyond the number crunching capability of even high-powered personal computers.

Fractal dimension (Fig 5.5) provides values that correlate with roughness measured in a self-similar way (Mandelbrot, 1982). The classic fractal problem asks what is the true length of the coastline of Great Britain. The answer is dependent on the scale of measurement. A satellite map, used to measure the coast in kilometres, will give a smaller value than that obtained if metre sticks were laid end to end along the high tide mark on the shore. This in turn would be smaller than the value obtained if a magnifying glass were used to measure along the edge of each grain of sand along the high tide mark, *etc.* A log plot of the length against the scale of the measuring instrument used shows that the fractal dimension in each case is self-similar. Obtaining this value of roughness is actually rather simple and straightforward. However, not many image processing software packages measure it — IMAGEPLUS and PRISM among few others do. As with other shape factors, fractal dimension is an attribute-specific rather than whole-shape measurement.

Topology measures the frequency of specific features of an object, such as the number of end points, loops, nodes and intersections (Fig. 5.6). This is done on the "skeleton" of an object; that is, the interior centre-line form of an object created by the computer by eroding the outside rows of pixels around the perimeter until the centre line (or network) of the object is reached. The numbers of points, loops and nodes are then counted and such values as the distribution of internodal distances can be calculated.

Combinations and permutations of the many ways to quantity attributes of size, shape, texture and topology provide the ability to address classification problems at the level of individual variation. Even if no two individuals are ever

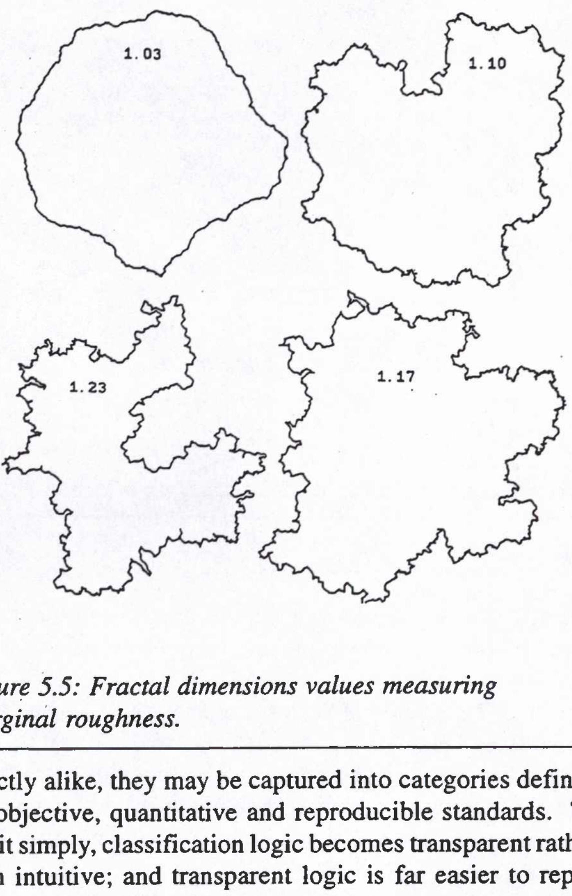


Figure 5.5: Fractal dimensions values measuring marginal roughness.

exactly alike, they may be captured into categories defined by objective, quantitative and reproducible standards. To put it simply, classification logic becomes transparent rather than intuitive; and transparent logic is far easier to replicate than is intuitive logic.

To date, I have applied IMAGEPLUS to a randomised spectrum of archaeological and fossil populations using rather simple and straightforward procedures. These tests have yet to challenge the capability of the system to resolve complex morphometric problems as occur in materials science or biomedical research. Nevertheless, even at simple levels of application the results can be overwhelming.

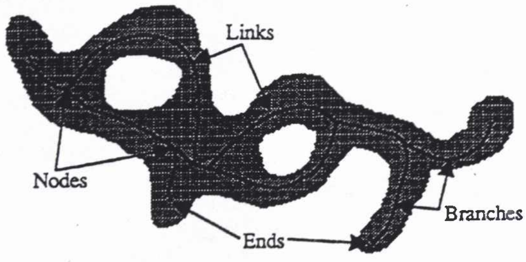


Figure 5.6: Topological measurement of central "skeleton" or "skeletal network" of an irregular object, allowing frequency counts and measurements of nodes, links, ends and branches.

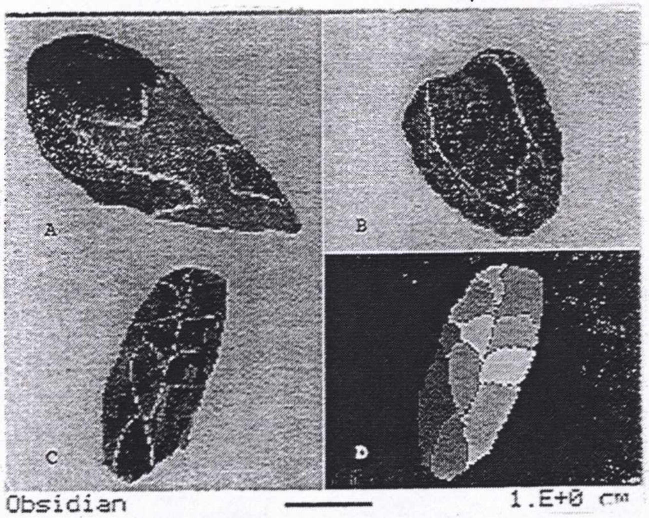


Figure 5.7: Isolating aspects of lithic artefacts for measurement of size and shape. (a) Isolating area of cortex remnant on a handaxe; (b) Isolating area of a Levallois flake scar on a core; (c) Isolating individual flake scars on a biface; and, (d) Grey scale level differentiation of individual flake scars on biface in (c).

5.4. Case 1: Looking at lithics

The ability to discriminate pieces and parts of objects allows for some potentially useful, unconventional information. With respect to Stone Age artefacts, for instance, it is easy to differentiate between secondary flaking on a large cortex flake (Fig. 5.7a) and area or cortex remnant. On a Levallois-style core, (Figs 5.7b and 5.8) the Levallois flake scar can be isolated to measure the percent of surface area of the core removed by the flake, as well as size and shape of the core and the flake. For a narrow biface (Figs. 5.7c–d), the eleven major flake scars were isolated to allow size and shape measurements of each of them simultaneously.

A slightly more involved test compared an isolated cache of 36 bifacial preforms with populations of standard projectile points from the same region. The computer was used to draw convex perimeters around the projectile points to approximate size and shape of the respective preform (Fig. 5.9). These reconstructions were then measured and compared by analysis of variance with the cache of bifaces. This unconventional, yet very simple way of doing typological cross-checking helped identify which projectile point types were potential end-products and which were not.

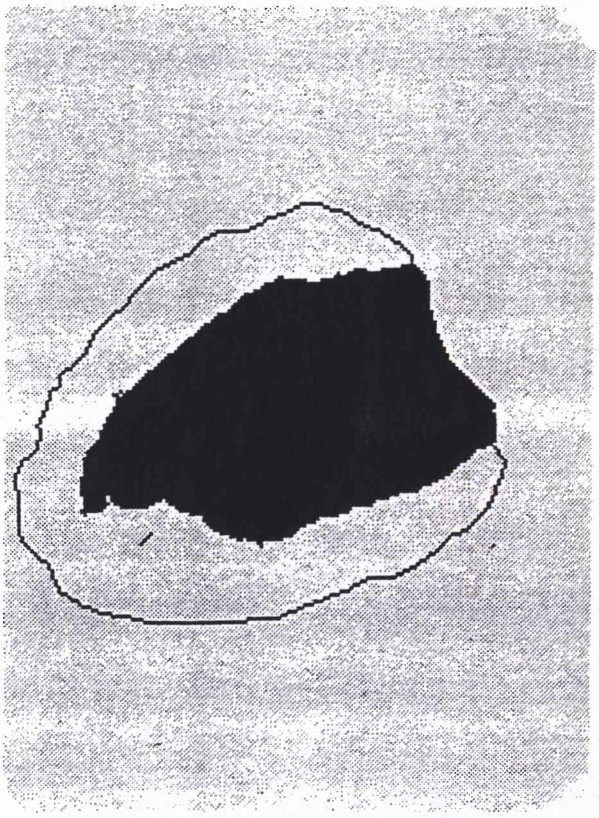


Figure 5.8: Isolating the major flake scar on a Levallois core for measurement of core and flake. Core area is 47.29 square centimetres; flake carried away 53.39% of core surface area.

Conventional length, width and thickness are hopelessly flaccid parameters inadequate to this kind of analysis. IMAGEPLUS is clearly an enhanced method for seriation and cross dating in archaeological analysis.

5.5. Case 2: Drawing a bead on technology

Dr. Geraldina Santini, Institute of Asian Studies (Naples) faced the problem of measuring thousands of small (typically 4–6mm in diameter) drilled stone beads, recovered from a Bronze Age cemetery in Oman. In less than two weeks, images of more than 2,000 beads were recorded for measuring. Subsequently over 100,000 measurements of size and shape were derived on bead top, bottom and side views, on bead attributes with the hole "filled", and on the drilled holes alone.

Preliminary statistical analysis presented an intriguing morphological pattern. When grave populations in the n = 30 to 90 range were plotted for distribution of aspect ratio values, measures which ignore the variable of drilled hole, the result was totally unexpected (Fig. 5.10). Instead of a normal, bell-shaped curve or minor departure from normal, the result was multimodal in the extreme — a plot of disconnected spikes representing 6 or 8 populations of up to 15 beads each. The obvious explanation is that the computer is reassembling beads with their neighbours cut from individual stone cylinders. If correct, it indicates that the integrity of the manufactured sets was maintained from the point of production to the point of deposition. This

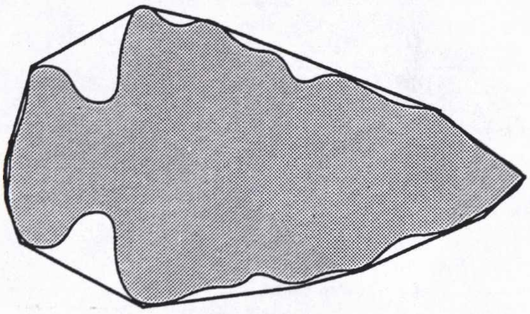


Figure 5.9: Construction of convex perimeter and area of a projectile point to approximate size and shape of preform allowing typological comparison with unfinished examples and workshop rejects.

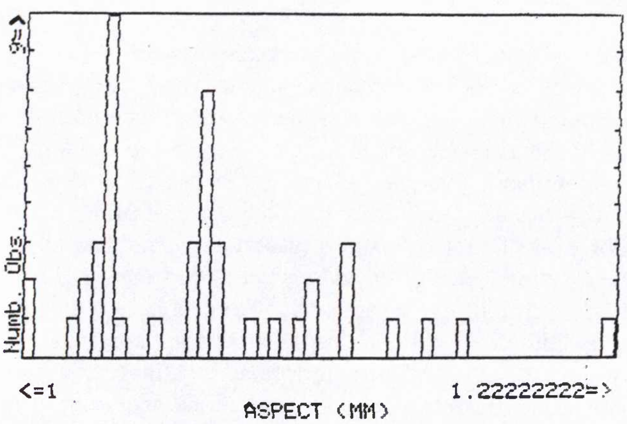


Figure 5.10: Distribution of aspect ratio values of drilled stone beads showing pronounced shape clustering, possibly an indication that bead groups cut from individual cylindrical cores were maintained as sets.

MIX #1 Blind Assemblage (n = 14)

Variety	*Actual*	*Area/perim*	*Added Parameter(s)*
Howding	4	4	not used
Sugar Pie	2	2	not used
Butternut	2	2	
Butternut or Pattypan		4	not used
Pattypan	4	0	
Acorn	2	2	not used

Pumpkin vs. Squash: 14 of 14 correctly classified _ 100%

Single Variety: 10 of 14 correct as assigned

Remainder: 4 correctly assigned to a double group

MIX #2 Blind Assemblage (n= 15)

Variety	*Actual*	*Area/perim*	*Added Parameter(s)*
Howding	2	2	not used
Sugar Pie	5	4	yes
Butternut	3	3	yes
Butternut or Pattypan		3	yes
Pattypan	3	0	
Acorn	2	2	not used

Pumpkin vs. Squash: 14 of 15 correctly classified _ 93%

Single Variety: 11 of 15 correct as assigned

Remainder: 3 correctly assigned to a double group

Error: 1 sugar pie pumpkin classed as squash.

TOTAL SCORE: Pumpkin vs. Squash: 28 of 29 correct = 96%

Subspecies assignment: 21 of 29 to single variety; 7 of 29 to two varieties; 1 incorrect

Table 5.1: Results of object recognition and identification of Cucurbita *seeds at the subspecies level.*

Species	*Area*	*Length*	*A'Ratio*	*R'Ndness*
T. monococcum	99.96%	91.77%	100.0%	100.0%
T. dicoccon	99.41%	100.0%	99.99%	99.99%
T. aestivum	80.39%	5.3%	19.32%	78.87%
H. vulgare	100.0%	100.0%	99.91%	99.96%

Table 5.2: Identification of "species unknown" wheat seeds (n = 14) recovered from an Iron Age mortuary vessel, using ANOVA probability of significant difference values on selected size and shape parameters.

leaves to the archaeologist the sublime problem of trying to figure out and assess the implications of the data.

5.6. Nature is not normal

I originally applied computer image analysis to deal with systematics of microscopic-sized plant opal phytolith fossils (Russ and Rovner, 1989) which I have been developing for more than 20 years (Rovner, 1971; Rovner 1983). Serious flaws in my own conventional classification logic were revealed along with other more general problems encountered using conventional typologies (Rovner and Russ, 1992). However, the most provocative results in archaeobotanical research to date were obtained recently in application of image analysis to identification and analysis of seeds. Although I had virtually no prior training or experience, the computer and I surpassed lifelong expertise in the identification of archaeological seeds literally within minutes.

The first test of seed identification using IMAGEPLUS followed a study by Decker and Wilson (1986) who used a somewhat simpler morphometric program to test seed populations of New World varieties of squash in three taxa, *Cucurbita pepo, Curcurbita pepo var. ovifera,* and *Cucurbita taxana.* Assignment of individual test seeds at essentially species level achieved an 86% success rate (Decker and Wilson, 1986, p. 601). I tested five reference populations at the subspecies levels, all seeds from varieties within *Cucurbita pepo.* Tests seeds were successfully assigned to squash versus pumpkin sub-categories with 96% accuracy (Table 5.1). More than 70% (21 of 29) of seeds in a blind test were correctly assigned to single cultivar varieties. Nearly 25% (7 of 29) were assigned to a double variety category, but only one of 29 assigned incorrectly. The earlier level of success was substantially surpassed.

A second opportunity to test seed identification arose during a resent research visit to the Archaeological Institute in Hungary. A population of "*Triticum*, species unknown" recovered from an Iron Age pottery vessel were digitised and measured using both size and shape parameters. Results were statistically compared, using analysis

		Imageplus			Schermann Atlas	
Species	(n)	length (mm)	mean	aspect ratio	length	aspect ratio
Cornus mas	76	-9.6914	12.1	**1.50-2.50**	11-16	**2.2-3.0**
Gingko loba	4	21.18-21.90	21.4	1.44-1.72	16-20	1.2-1.4
Lupinus polyphyllus	106	-3.605	**4.6**	1.10-1.90	**3.8-4.8**	1.5-1.8
Malus germanica	30	10.4-13.3	**11.9**	**1.18-1.55**	**8-12**	**1.8**
Prunus spinosa	21	11.8-14.3	13.1	1.20-1.54	7-9	1.5-1.6
Prunus spinosa	14	-8.261	9.1	1.09-1.33		
Sambucus nigra	75	2.81-4.56	3.8	**1.38-2.40**	3.5-4.5	1.8-2.2
Sambucus racemosa	79	3.47-4.94	**4.1**	**1.09-1.65**	**2.2-3.2**	**1.8-2.2**
Solanum dulcimara	132	1.98-2.97	**2.5**	**1.00-1.56**	**2.2-2.6**	**1.5**
Sorbus acuparia	89	2.97-4.95	**4.4**	**1.71-2.55**	**3.4-4.8**	**3.0**
Sorbus domestica	28	6.60-8.75	**7.7**	1.28-1.81	**6-7**	1.4-1.5
Tilia platyphyllos	88	6.46-12.56	8.6	1.00-1.75	7-10	
Viburnum lantana	19	6.44-9.91	**8.1**	1.14-1.71	**6-7**	1.1

Table 5.3: Computer morphometry vs. reference atlas data (Scherman 1966). Comparitive values which disagree substantially are shown in bold type.

Species	(n)	Skew	Kurtosis	Normal?
Cornus mas	76	-.182	2.219	no
Lupinus polyphyllus	106	.208	3.024	maybe
Sambucus nigra	75	-.989	4.776	no!
Sambucus racemosa	79	.366	2.467	no
Solanum dulcamara	132	-.010	2.589	maybe
Sorbus acuparia	89	-1.30	6.842	no!
Tilia platyphllos	88	1.04	5.650	no!

Table 5.4: Test of normal distribution of length in seed populations.

of variance (ANOVA, f-test), to previously measured reference seed populations of 4 cereal grain taxa (Table 5.2). The results, rather unequivocally, assigned the unknown to the species, *T. aestivum*, bread wheat — an inexperienced computer operator surpassing lifelong expertise with less then 15 minutes of effort.

As impressive as such results in seed identification may appear, they are, in fact, misleading and subject to verification. ANOVA is a parametric test which assumes a normal distribution. In further testing of seed populations, we soon noted that histograms of seed size distribution were plotting strangely in spite of the fact that we were measuring rather large populations of seeds. Within minutes we tested enough seed populations to indicate, in compelling fashion, that seed size distributions are not bell-shaped or normal (Table 5.3). Parametric tests of seed size are therefore inappropriate and any previous statistical studies of seed identification which employed parametric statistical tests are suspect.

More morphometric heresies appeared when computer generated measurements of large seed populations were compared against data presented in standard reference seed identification atlases. Comparison with the Schermann (1966) atlas were particularly relevant because we were using the same seed reference collections from the Hungarian Museum of Agriculture. In a majority of cases tested and compared, manual measurements of small seed samples — usually a population of ten seeds (Montgomery, 1978) — were wholly unreliable, badly misrepresenting individual seed variation in a population (Table 5.4). The mean size value of several computer-measured populations fell near or beyond, sometimes well beyond, the minimum or maximum range value in the reference manual. Since archaeological seed populations are far more likely to be random representations of large reference populations, size categories are therefore very likely to be misinterpreted. For example, interpretations of climatic conditions or domestication of taxa based on large seed size values are clearly suspect if comparisons are made with data in such standard seed atlases. If parametric statistical tests are employed, the level of suspicion increases. Clearly, a new, morphometric reference atlas based on size and shape measurements of far more than ten seeds in a populations is essential for archaeobotanic research.

Seeds are only the beginning. There is every reason to expect that nature is not normal in a host of other botanical and zoological systems used in archaeological research.

5.7. Conclusion

Computer image analysis and stereological morphometry should substantially replace the simpler paradigms of quantitative methods in critical areas of archaeological research. The normative type is itself typically an unreal and arbitrary stereotype, an unrepresentative fraud which should be relegated to the dustbin without regret. Likewise, it is time to stop worshipping the normative mean value and the Gaussian bell-shaped distribution curve. These icons are no longer worthy of *unquestioned* adoration in the morphological analysis of archaeological assemblages and fossil populations. The age of non-parametric statistical analysis may be at hand.

Bibliography

DECKER, D. & H. D. WILSON 1986. "Numerical analysis of seed morphology in *cucurbita pepo*", *Systematic Botany* 11(4): 595–607.

MANDELBROT, B. 1982. *The fractal geometry of nature*. W. H. Freeman, San Francisco.

MONTGOMERY, F. H. 1977. *Seeds and fruits of plants of eastern Canada and northeastern United States*. University of Toronto Press, Toronto and Buffalo.

ROVNER, I. 1971. "Potential of opal phytoliths for use in paleoecological reconstruction", *Quarterly Research* 1(3).

ROVNER, I. 1983. "Major advances in archaeobotany: Archaeological uses of opal phytolith analysis", *in* Schiffer, M. (ed.), *Advances in Archaeological Method and Theory* 6. Academic Press, New York.

ROVNER, I. & J. C. RUSS 1992. "Darwin and design in phytolith systematics: Morphometric methods for mitigating redundancy", *in* Mulholland, S. C. & G. R. Rapp (eds), *Phytolith systematics: Emerging issues*. Plenum Press, New York.

RUSS, J. C. 1986. *Practical stereology*. Plenum Press, New York.

RUSS, J. C. 1992. *Computer-assisted microscopy: The measurement and analysis of images*. Plenum Press, London.

RUSS, J. C. & I. ROVNER 1989. "Stereological identification of opal phytolith populations from wild and cultivated *Zea mays*", *American Antiquity*, 54(3): 784–792.

SCHERMANN, S. Z. 1966. *A magismeret atlasza*. Mezögazdaság Könyvkiadó, Budapest.

Dr Irwin Rovner
North Carolina State University
Box 8107
Raleigh, NC 27695-8107
USA
IRV@server.sasw.ncsu.edu