Bookmarks Set for Section & Figure Views

Features of Derivative Continuity in Shape

PHILIP F. HENSHAW
~~167 W 87th St New York NY 10024~~
for publication IJPRAI 12/99
id@synapse9.com

Abstract

Derivative continuity is a distributed invariant relationship between parts of flowing shapes. The original techniques presented here were developed for making the behavioral dynamics of complex processes more recognizable, but are equally applicable to assisting in the recognition of shapes in images. Regularizing a sequence using a constraint of derivative continuity is equivalent to using a bimodal smoothing kernel, producing a distinct bias for reducing variation on higher derivative levels, sharply defining shape with minimal suppression of shape. To help determine where reconstructing shapes this way is valid a test was developed to help distinguish combinations of noise and smooth shapes from random walks. This helps distinguish between illusory and genuine data shapes, but also exposes a flaw in using this and other measures of scaling behavior for diagnostic purposes. Gaussian scale space techniques in use for some time for identifying reliable landmarks in the shapes of outlines, are demonstrated for use in identifying key features of shape in time series.

Key Words: PATTERN RECOGNITION, DERIVATIVE CONTINUITY, REGULARIZATION, RECONSTRUCTION, SMOOTHING KERNELS, CONNECTILES, FRACTAL EXPONENT, SCALE SPACE .

1 Introduction

In a literal sense, equations that represent physical behaviors are simplifications of the shapes of the data gathered about those behaviors, converted to a more useful form. That form may be more useful because of its compactness and ease of manipulation, but also frequently for seeming to embody the very structure of the behavior whose shape it imitates. The methods reported here concern one of the invariant relationships between near bye points in equations with smooth shape at the core of the special relationship between equations and nature, derivative continuity. The property is used here, , to regulate non-parametric representations of shapes with the object of better representing underlying behavior and its differential structures. That may be preferred over the use of equations for the purpose in preliminary investigations, when the subjects of study are uniquely individual or transient, or for identifying natural markers of shape or behavioral structure for object or process recognition.

Part of the underlying effort was to define continuous derivatives for sequences, using the mathematical definition of the derivative and parsimony to construct smooth curves passing through given points by iteration. For mathematical functions, having a slope at a point, a derivative, is determined by whether the slopes, between a given point and points on the curve approaching it from opposite directions, come to the same value. This definition and test for a derivative has been well worked out for mathematical functions for a long time, originally developed by Newton and Leibnetz in the invention of calculus[4]. Stated simply, curves which have a slope at every point change by continuous progressions in direction rather than discontinuous steps in magnitude, and represent distributed structures of connection.

Derivative continuity, interestingly, is not defined for the smooth flow of physical processes, except in equations representing them, nor for smooth shapes in data. For physical shapes and processes the connectedness of smooth flows is a presumption, one that has proved to be

singularly useful in allowing the representation of physical behavior with equations. There are, of course, disjointed, random, fractile and chaotic patterns that seem common as well. In fact all things and processes are so full of gaps as to seem perfectly discontinuous in ultimate construction. Still the great success of presuming continuity, provides direct support for considering it as an invariant physical principle of meaningful shape, even if not well understood. Except for behaviors related to the physics of the conservation laws, the physical basis for derivative continuity in nature is relatively unstudied.. It is this study of the physical basis of continuity that has motivated much of this work. Progress on that wider interest is fairly limited but it has produced some techniques that should be independently useful.

1.1 Related Issues and Methods

The object of generalizing data shapes is to create features that distinguish the various subjects represented in scattered points of data. This requires applying constraints to what would otherwise be 'ill-posed' problems, applying various assumptions regarding the non-accidental character of nature as discussed by Wechsler[17] in relation to the problems of computer vision and artificial intelligence. Derivative continuity is typically part of the constraint applied in the use of Lagrange multipliers[17], for example, as well as in fitting 'least energy' surfaces such as with the level set approach[18] , because the constraints applied are continuous functions. Continuity can also be applied using constraints on local connections between points that are weaker than optimized equations, with the benefit of preserving subtle and transient characteristics of shape that would otherwise be lost.

To determine whether it is appropriate to assume underlying continuity and the validity of apparent statistical trends or shapes, variation scaling and range tests[3,5,11] have been used. These same tests underlie the measures of fractile dimension for natural scaling patterns and have been used to identify a wide range of apparent fractile patterns in nature. One of the invariants of natural data, however, is a presence of mixed signals, originating from mixed sources. That all variation scaling and range tests are compromised by sensitivity to mixed signals has been noted[16] but is not widely investigated. Because some the measures display differing sensitivity to different data structures, however, better ways to separately identify components of mixed types of data may be developed.

Once meaningful data shapes are found it is common to use non-linear operators to isolate identifiable landmarks[8]. The Gaussian scale space methods[7,12,13,15] , in simple terms, use inflection points on a curve that are relatively invariant to smoothing as identifying marks of shape. In concept, though not yet in practice, the location and kind of derivative inflection points on a smooth shape can be read like a bar-code, uniquely identifying given subjects. Given sufficiently accurate representation of shapes, various other details of the general differential structure can also be tested for robustness and used as identifying marks. While most work done in the area concerns the outline shapes of objects in visual images the method is also directly adaptable to identifying scales and landmarks of differential structure in time series.

2 Means and Methods

2.1 Derivative interpolation and smoothing

A smoothing 'kernel' is a distribution of factors with a sum of one, such as:

$$\text{for} \quad -a < n < a \qquad k(n) = \frac{e^{-|n|/2}}{\sum_{n} e^{-|n|/2}} \tag{1}$$

giving a Gaussian distribution of weights to group of points, used to calculate contributions from neighboring points in smoothing a sequence by :

$$U^1(Y,k) = \sum_{i=0}^{j} \left| \sum_{n=-a}^{+a} k(n) \cdot Y_{i+n} \right. \tag{2}$$

Here **a** is the number of points on either side of each point **i** in sequence **Y** of length **j**. The kernel **k** will be undefined for values of **i** near **0** and **j** unless altered in those regions. Repeated smoothing of sequence **Y** by kernel **k** is indicated $\mathbf{U^m()}$.

Gaussian shaped kernels (figure 1, $\mathbf{k_1}$) are considered to be ideal for scale space techniques [7, 14] particularly for their broad invariance to transformation[15]. The kernel $\mathbf{k_2}$ shows linear center weighting for comparison. These definitions may be applied to any size kernel, indicated $\mathbf{k_{1,15}}$ for example, for Gaussian smoothing using a range of 15 points.

A bimodal smoothing kernel (figure 2, $\mathbf{k_3}$, $\mathbf{k_{3R}}$) results from an algorithm derived for derivative smoothing. A bimodal kernel suppresses the contribution of the center point and enhances the contributions of the adjacent points. The effect is to enhance distributed relationships between points and suppress the effect of individual point values.

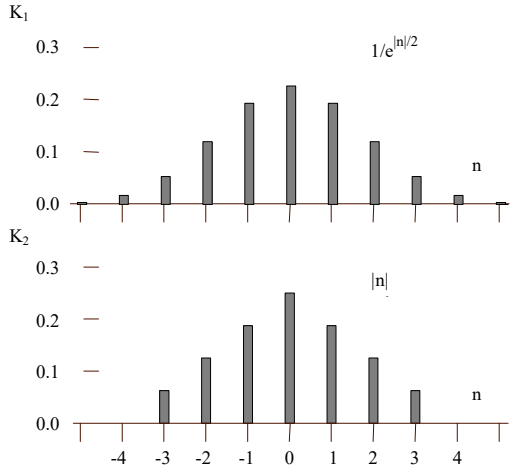


Figure 1. *Normal Gausian and linear center weighted smoothing kernels.* $K_1=1/e^n / \Sigma_m 1/e^n$, $K_2=(m-|n|)/\Sigma_m (m-|n|)$ shown for m=5 and 4 respectively. Dividing by the sum of weights assures that the sum of the factors will equal 1.

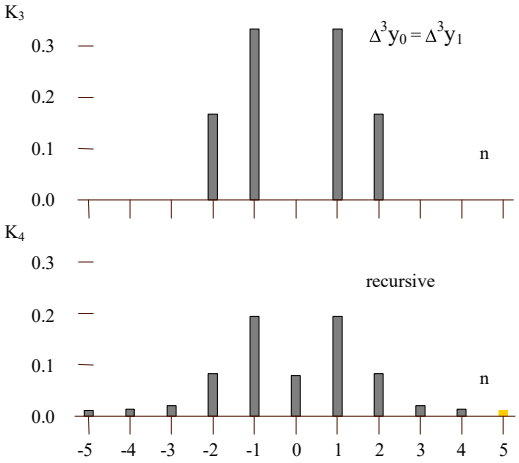


Figure 2. *Derivative smoothing kernels*. Equalizing third differences at a point using two points on either side (figure 3). K_{3r} results from using the values produced by K_3 in place of the originals in the sequence as each successive point is evaluated. The sum of values shown for $K_3 = 1$, for $K_{3r} = .985$

The derivation of the simple bimodal kernel ($\mathbf{k_3}$) follows directly from applying a condition approximating derivative continuity as a principle of connection between five points. The algorithm calculates the value needed to equalize the third differences for the first and last groups

of four points (approaching from either side) in a five point set (figure 3). This is equivalent to calculating a point on a curve passing through the adjacent points for which the fourth derivative is zero, a 3rd degree polynomial. This is applied recursively to produce the kernel $\mathbf{k_{3R}}$, using each adjusted point in place of the original value in calculating the next. This overlapped linking of five point connectiles produces a distributed continuity and spreads the weights applied to each point throughout the curve. It effects the smoothness of a sequence more than its shape, and produces curves very closely fitting the original data.

For the case of equally spaced points the value of the new point is:

$$y_2 = \frac{2}{3}(y_1 + y_3) - \frac{1}{6}(y_0+y_4) \tag{3}$$

This relation is arrived at by first taking the successive differences with $\mathbf{y_2}$ as the unknown:

$$\Delta x_0=x_1-x_0=1 \qquad \Delta y_0=y_1-y_0$$
$$\Delta x_1=x_2-x_1=1 \qquad \Delta y_1=y_2-y_1 \qquad \Delta^2 y_0=y_2-2y_1+y_0 \qquad \Delta^3 y_0=y_3-3y_2+3y_1-y_0 \tag{4a}$$
$$\Delta x_2=x_3-x_2=1 \qquad \Delta y_2=y_3-y_2 \qquad \Delta^2 y_0=y_3-2y_2+y_1 \qquad \Delta^3 y_1=y_4-3y_3+3y_2-y_1 \tag{4b}$$
$$\Delta x_3=x_4-x_3=1 \qquad \Delta y_3=y_4-y_3 \qquad \Delta^2 y_0=y_4-2y_3+y_2$$

and then setting equation 4a equal to 4b.

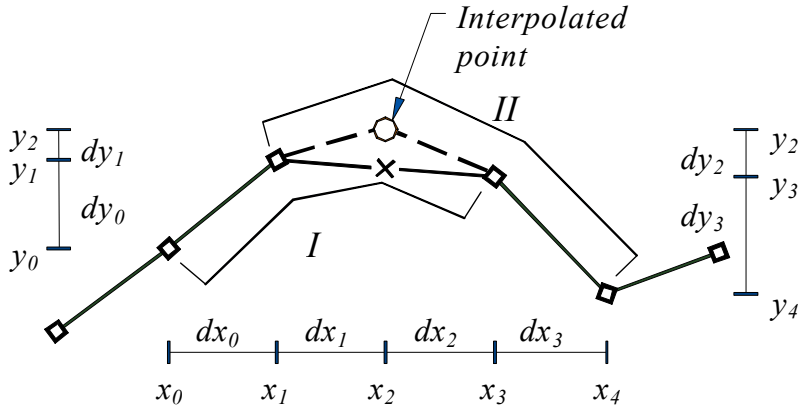


Figure 3. *Interpolating a Point to Approximate Derivative Continuity*. To locate the middle point, lying on a 3rd degree polynomial through four points, the value y2 is calculated to make the 3rd differences in periods I and II equal.

Because the algorithm does not make use of the value of the point in question it can both be used for smoothing and interpolation, inserting a new point in the middle of a four point group. In figure 3 the point 'O' is inserted on the path of the implied 3rd degree polynomial through the other points. The point 'X' is the point that would result from linear (1st derivative) interpolation. The scheme for 4th or higher derivative interpolation, or for using non-polynomial connectiles, would follow the same approach, though these have not been tested. When used for repeated smoothing, the recursive kernel ($\mathbf{k_{3R}}$) rapidly reduces fluctuation in the higher derivatives with rapidly declining effect on the path of the curve.

There are also a few biases and distortions that can develop with this approach, and reasonably successful corrections made for them in the implementation. Using recursive smoothing produces a directional bias which is compensated for by separately scanning the sequence forward and back and averaging. The algorithm is undefined near the ends of a sequence for lack of a sufficient number of points. In most cases dropping these points would degrade the data more rapidly than modifying the routine to use fewer points or using extrapolations for the missing points. The preferred choice has been to use damped 2nd derivative projections for the missing points and to reduce the confidence in interpreting shapes near the ends of the resulting curves.

Sometimes the point that would minimize the 3rd derivative at a point introduces a reversal in sign of a 2nd derivative within the group of points. Sometimes this is exactly what continuity implies, and sometimes it exaggerates coincidental shapes inappropriately. Simply cushioning this effect has proved reasonably effective. Highly irregular point spacing complicates the analysis, and requires making additional assumptions, for any kind of smoothing routine. The present routine makes a correction for moderate irregularity in point spacing and curves with highly irregular point spacing are given regular point spacing by periodic integration before their shapes are studied.

A continuous curve is constructed from a finite sequence using $\mathbf{k_3}$ and $\mathbf{k_{3R}}$ in alternation, first for interpolation and then smoothing. The result approximates an infinite degree polynomial. Given the qualifications noted, repeated derivative interpolation and smoothing provide a reliable iterative method of defining a slope at a point for finite sequences.

2.2 Difference scaling exponent, H_σ

Accumulations of random events tend to produce shapes giving false impressions of the processes producing them. The need to rule out the presence of random walk to be able to define slopes in data has been given special attention as in paleontology[3,16] concerning one of the data sets to be used in the demonstration section below. When the nature of the underlying processes are unknown, implications of the apparent shapes unclear, and random walk mechanisms conceptually plausible, there may be no basis for interpreting shape without statistical tests to measure accumulations of random change. Such sequences have randomly varying rates of change (random first differences) so that no rate of change has predictive value for any other, an absence of meaningful shape..

A random walk, **R(n)**, is a sum of independent copies of a random variable:

$$R(n) = \sum_{n} X \tag{5}$$

where **X** is a random variable with variance of ν, and the mean end point variance for all such **R(n)** increases as $n \cdot \nu$. Consequently the variance of the differences over shorter intervals will be expected to increase in proportion to the length of the interval. The set of step differences in **R(n)** of interval length **k** is:

$$\Delta R_k = \left(\sum_{j=0}^{k} X_j \, , \sum_{j=1}^{k+1} X_j \ldots \sum_{j=n-k}^{n} X_j \right) \tag{6}$$

composed of $n-(k-1)$ sums of **k** step differences, with expected mean variance of $k \cdot \nu$.

A sequence combining ordinary Gaussian noise with another behavior:

$$S(n) = f(n) + X_n \tag{7}$$

where $f(n)$ is unknown, has single step differences of $(f(n)+X_n)-(f(n-1)+X_{n-1})$ and step differences of interval length **k**:

$$_{\Delta}S_k = \left(f(k) - f(0) + (X_k - X_0),\ \ldots f(n) - f(n-k) + (X_n - X_{n-k})\right) \tag{6}$$

providing $n-(k-1)$ step differences, with a mean step variance of $2\nu + \nu_f$, where $_f$ is the variance for step differences in the unknown. The step variance of the noise component, 2 , will be the same for all interval sizes, and the combined variance will depend on the structure and

component approximates a simple straight line trend at the scale of the intervals considered its step

This effect is seen in the slopes of a log/log plot of **ν** and **k**, allowing a measure **ν'** to be defined as that slope:

$$\nu' \cdot \ln\frac{k}{j} = \ln\frac{\nu(_{\Delta}S_k)}{\nu(_{\Delta}S_j)} \quad \ldots\ldots \tag{8}$$

where **j** and **k** are the step sizes used to make sequences $\Delta\mathbf{S_j}$ and $\Delta\mathbf{S_k}$ of the step differences of **S(n)**, with $\nu(\Delta\mathbf{S_j})$ and $\nu(\Delta\mathbf{S_k})$ their variances, where the expected range of **2>ν'>0** with **ν'=1** for random walks of a Gaussian random variable. Stated in the conventional form for difference scaling in fractiles,

$$\sigma(\Delta_k)=ce^{H\sigma} \tag{9}$$

replacing **ν** with its square root **σ** and **ν'** withexponent $\mathbf{H_\sigma}$ having an expected range $\mathbf{1>H_\sigma>0}$ with mean of $\mathbf{H_\sigma = .5}$ for pure random walks, as for the Hurst exponent ($\mathbf{H_\Delta}$) used to measure fractile dimension[5,11].

$$E(|\Delta_k|)=ce^{H_\Delta} \tag{10}$$

For standard comparison the value of **c** is to be taken as the value of the measure on the minimum length interval, so that the initial value of the log/log plot is always one (1) as seen in figures 7, 8, 9 & 11.

3 Demonstrations

Brief examples are offered to illustrate these techniques. The derivative reconstruction and Gaussian smoothing work was done on a P.C. in AutoCAD by Autodesk using its lisp programming language and are available as a series of lisp routines for AutoCAD v12-13 from http://idt.net/~ph/Curve.zip. Any general analysis platform allowing varying point spacing for different measures in a single table would be serve as well. The step variance and fractile measure work was done in JMPIN by SAS Institute and is available in JMP format from http:// idt.net/~ph/StepVar.zip.

3.1 Fine Shapes of Gamma Ray Burst BATSE Trigger 551

Figure 4. shows a sequence of 150 rates of gamma ray detection at BATSE satellite detectors[1] including a gamma ray burst ($\mathbf{Y_1}$). The raw measure was composed of 191 points with irregular spacing, based on the length of time between every sixth photon recorded. This irregularly spaced sequence was converted to the regularly spaced 150 point sequence shown by periodic integration over .5 millisecond intervals, calculating the areas under the trace of the original points. In figure 5 the curve $\mathbf{U^1(Y_1,k_{1\text{-}11},)}$ is the result of a single pass Gaussian smoothing, combining each point with five points on either side. The other curve, $\mathbf{U^4(Y_1, k_3,)}$ shows the effect of four steps of derivative smoothing, first applying $\mathbf{k_{3R}}$ for smoothing, then interpolation with $\mathbf{k_3}$ to make a curve with 300 points, and then smoothing twice with $\mathbf{k_{3R}}$, the basic steps of derivative reconstruction. The naming convention indicates four repetitions of $\mathbf{k_3}$ for convenience in notation.

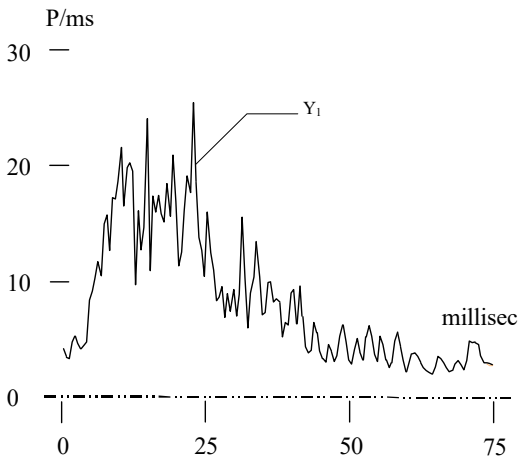


Figure 4. *Gamma ray burst BATSE 551.* Photons per millisecond, source NASA, Burst And Transient Source Experiment. Time of each photon in microseconds, from detectors 0, 2 & 6, energy levels 3 & 4, raw rate for each sixth photon leaving191 points. The 0.5 milliseconds averages leave 150 regularly spaced points for uniform smoothing.

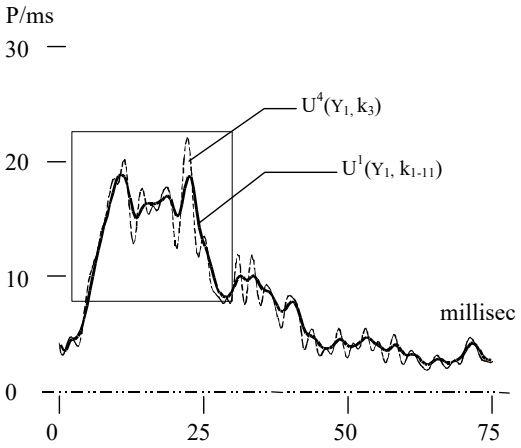


Figure 5. *Center weighted averaging and derivative smoothing*. For gamma ray burst data of figure 4, using one pass of Gaussian smoothing $U^1(Y_1, k_{1\text{-}5})$ and four passes of recursive smoothing $U^4(Y_1, k_4)$, for both smoothing (S) and interpolation (I), in the sequence S-I-S-S. Detail box shows area enlarged in figure 5

The differing shape suppression effect of the two methods is clear, but the real question is whether the more detailed shapes provided by derivative reconstruction are real features of the data. This is demonstrated in figure 6 showing an enlargement of the detail box of figure 5. The same curves are shown along with five others made by the same steps of derivative reconstruction based on six independent subsets of the original data. The six subsets were made by the same method, selecting every sixth point, but beginning from different starting points in the original sequence. The six curves visibly follow each other very closely and show tightly clustered turning points as

[1] The BATSE Trigger 551 photon arrival data contains a list of 28904 photons. The data shown is from detectors 0,2&6 in energy levels 3&4, selected for least noise, showing the third of the three gamma ray bursts in the particular record, representing 1141 photons.

indicated by the close groups of vertical lines which mark the minima of all six curves. Virtually every little bump turns out to be highly robust to resampling the original data, indicating that virtually none of the shapes originate from either the smoothing method used or from chance sampling of the data. The scientific question of whether these shapes represent any particular physical process or event is, of course, a separate question and some might represent illusory shapes of Poisson noise as well as blurred images of far off explosive events. The comparison also shows how Gaussian and derivative smoothing are completely consistent with each other, representing differing degrees of generalization of the same robust underlying shapes, many of which are not at all clearly evident from visual inspection.

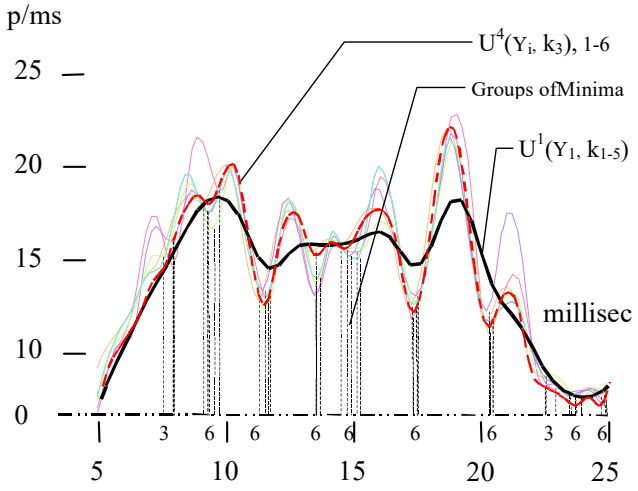


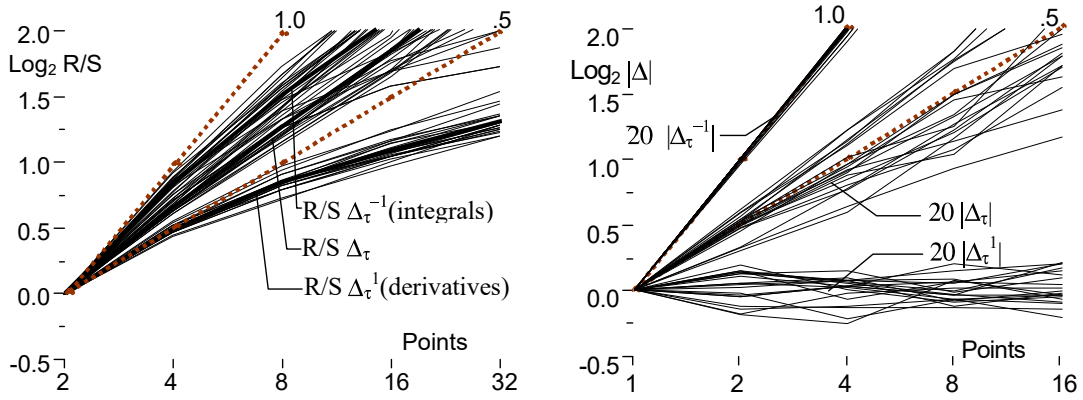
Figure 6. *Detail of figure 5 with overlay of derivative smoothing for 5 other independent subsets* Each subset was formed from sampling one of six points from the BATSE 551 data from different starting points, ~75 points each. The numbers of roughly coincident minima are shown below the axis. The derivative smoothing shown in figure 5 is the dashed line.

3.2 Comparing Measures of Scaling

There has been considerable interest in random walk patterns in relation to 'Hurst's Law', relating power law scaling of variation to long range dependence within natural sequences, further developed by Mandelbrot to characterize fractile patterns[10]. Equation 10 is the basic measure presently used to determine fractile dimension [5,11]. The rather close relationship of fractile scaling measures to the step variance test for symmetric noise, equation 9, was found during final literature search. The established measures of scaling behavior correlate the means of change over intervals of different sizes, and the new measure correlates their standard deviations. More significant differences lie in the way the tests are intended to be used. Figure 7a and 7b show Hurst's original correlation, $\mathbf{H_{R/S}}$, and Mandelbrot's, $\mathbf{H_{\Delta}}$ for the same 20 arbitrarily chosen random walks of 96 points. Figure 8a shows $\mathbf{H_{\sigma}}$ applied to the same random walks, and 8b their traces along with the $\pm\sqrt{2n}$ expected range for all random walks.

The rationale for examining the variance (or standard deviation) of stepes in a sequence at different scale intervals was to see if the internal variation in successively larger periods would cancel out. That would indicate the presence of noise. If the variance of change is relatively constant for different size intervals it indicates that adding successive points cancels out the

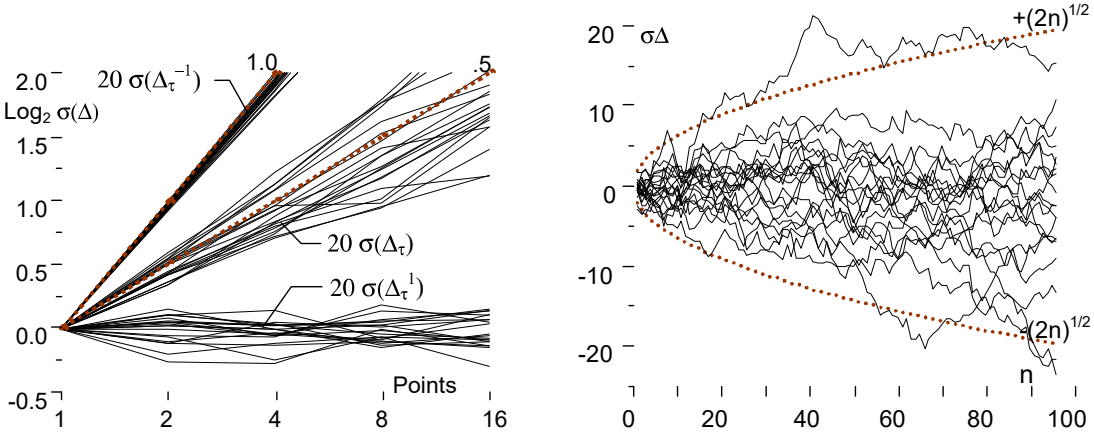
variation of preceding points, indicating symmetric variation about some norm. This is preliminary evidence of underlying continuity, having a norm, and justifies noise reduction and an examination of the underlying shape for recognizable behavior.



7a) $H_{r/s}^{1}$=0.33, $H_{r/s}$=0.61, $H_{r/s}^{-1}$=0.63
σ=0.022 σ=0.062 σ=-0.11

7b) H_{Δ}^{1}=0.0041, H_{Δ}=0.48, H_{Δ}^{-1}=0.99
σ=0.031 σ=0.098 σ=-0.033

Figure 7. *Fractal scaling exponent H by traditional methods* a) Hurst R/S indicator, $H_{r/s}$: $E(Y_{\Delta max}-Y_{\Delta min})/E(\sigma(Y'_{\Delta}))=c_1 t_{\Delta}^{H}$, b) Mandelbrot absolute first differences measure, H_{Δ} : $E(|\Delta_{\tau}|)=c_2 * t_{\Delta}^{H}$, for twenty 96 point random walks; H for random walks, H^1 for their derivatives, and H^{-1} their integrals, using ratio of longest to shortest period, and giving standard deviations.



8a) H^{1}=0.0047, H=0.45, H^{-1}=0.90
σ=0.031 σ=0.088 σ=-0.053

8b) 20 random walks used for 4a, 4b, & 5a

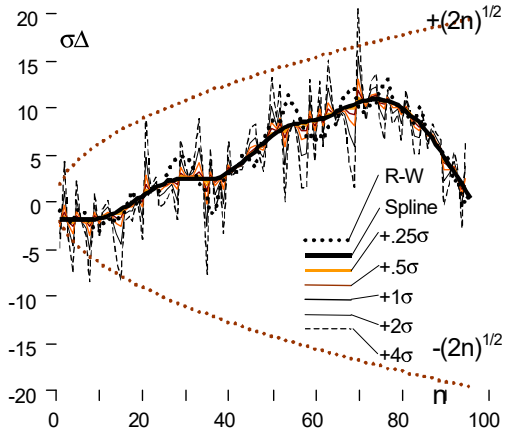
Figure 8. *Scaling exponent H for variance symmetry test.* a) Standard deviation of first differences, H_{σ}: $E(\sigma(\Delta_{\tau}))=c_2 * t_{\Delta}^{H}$ for 2,4,8 &16 point periods in 20 random walks of 96 points; b) Plots of the 20 random walks used

Figures 7 & 8a also show the performance of each measure for the derivatives and integrals of each of the 20 random walks as well as each of the random walks themselves. The differing behavior of derivatives and integrals can expose more of the structure of a sequence being studied. A number of experiments to distinguish the ratios of noise, random walk and continuous shapes were done with some small success, but it is a difficult subject and a general discussion is beyond

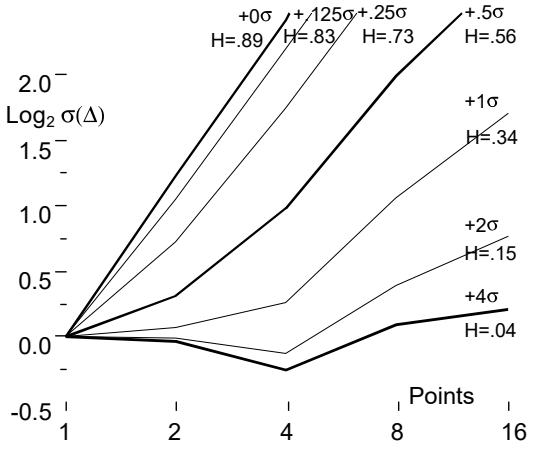
the present scope[6]. The high values of **H** for the integrals of random walks demonstrates the positive correlation of steps in their integrals, the same as the high values of **H** expected for continuous curves[10].

The typical pattern of effects due to added noise in a sequence is demonstrated in figure 9a and 9b. The standard used is slightly complex, but offers a nice demonstration of the strong and directly proportionate effect of noise in reducing the apparent scaling behavior of other components of the sequence. A random walk is used as the basis for creating a smooth shape using a 100λ spline. Then various degrees of noise, as fractions or multiples of the step variance that generated the original random walk, are added and the values of **σ(Δ)** and the slope H calculated for each (figure 9b). The slopes, and corresponding scaling exponent of the data are reduced in direct proportion to the degree of noise in the data, reducing the slope to near or below zero when noise dominates. The unfortunate consequence is that mixed signals composed of combinations of continuous and random walk curves and moderate noise may have any scaling exponent. Even simple quadratics, with just a little added noise, may display constant power law scaling of variation simulating random walks[6].

Sometimes it is possible to tease these components apart by comparing the scaling exponents of derivatives and integrals. Continuous shapes tend to have high scaling exponents (H=.89 for the spline curve in figure 9a) and tend to be unchanged by integration or differentiation, while the integrals or derivatives of random walks and noise are strongly effected, and tend to fall in different ranges[6]. Generally, though the problem is under-constrained and requires additional information or assumptions to solve.



9a) Spline of R-W +.125σ,.25σ,.5σ, 1σ, 2σ, 4σ



9b) H= 0.89, 0.83, 0.73, 0.56, 0.34, 0.15, 0.04

Figure 9. *Effect of Noise on Scaling Exponent H.* a) a random walk as in fig. 8b., with a smooth 100λ spline, to which various amounts of noise are added, in factors of σ for the steps of the given random walk. b) plots of σ(Δ) for each curve of 9a. showing regular decline of slope for increasing noise.

Comparing the performance for derivatives and integrals also displays differences between the behaviors of the measures themselves. Hurst's original formulation (fig. 7a) has complex scaling effects within the measure itself, using the range of movement **(R)** within each time period **(t_Δ)** rather than change over the whole period **(Δ_τ)**, and dividing by the standard deviation of step differences within each period without accounting for the changing standard error of that measure

for each interval period. It is clear from the figure that the measure does not clearly distinguish the behavior of the integrals, and has such narrow variation that it appears insensitive to changes in data structure in general. The values of $\mathbf{H_{r/s}}$ also regularly exceed those of $\mathbf{H_\Delta}$, failing to have a mean of .5 as expected, a discrepancy also noted by Hastings and Sugihara[5] (p 151). The behavior of the standard measure, $\mathbf{H_\Delta = E(|\Delta_\tau|)}$, appears all but identical for random walks to $\mathbf{H_\sigma = \sigma\,\Delta_\tau}$ seen in figures 7b and 8a. The two measures, however, provide significantly different results for other kinds of sequences (see figure 11). There are also a number of other methods for calculating values for the scaling exponent with a likelihood of differing behavior. While an effort has been made to be balanced and complete in relating the new measure to the preceding ones, a full discussion is beyond the present scope.

3.3 Evolutionary rates of G. tumida

G. *tumida* is a common photosynthesizing plankton, which apparently tripled in size in a rapid transient process around 5.5 million years ago, as studied in some detail by B. Malmgren[9]. The data shows 95 samples representing about 50 specimens each, spanning the past 6.5 million years, collected from an Indian ocean sediment core (figure 10). This data has been discussed as a possible random walk[3,16] because random walk can not be automatically ruled out for the progression of genetic variation. Unable to prove otherwise, it is considered possible that genetic structures could vary as freely as location of molecules in Brownian motion, even the common genetic structures of vast populations.

The relative step variances shown in figure 11, for 1, 2, 4, 8 and 16 point step intervals, however, indicate that non-accumulative variance dominates. The value of $\mathbf{H_\sigma}$ is .17, almost four standard deviations (see fig 8a) below the mean of pure random walks. This indicates that the structure of the variance is primarily symmetric and justifies appropriate noise reduction to see if recognizable behavioral shapes can be identified. The value of $\mathbf{H_\sigma^{-1}}$ is .89, showing an increase from $\mathbf{H_\sigma}$ that is far greater than would be expected if the sequence were purely composed of noise, suggesting that the noise at the integral level is less successful in masking the presence of an underlying process with strong serial correlation.

The values of $\mathbf{H_\Delta}$, using the $\mathbf{E(|\Delta|)}$ method, arealso shown in figure 11 for record and comparison. The significant differences appear to be that unlike the value for $\mathbf{H_\sigma}$, the value of $\mathbf{H_\Delta}$ for the integral of the data, $\mathbf{H_\Delta^{-1}}$, is at the limit of the expected range, 1.01, and unlike the value for $\mathbf{H_\sigma}$, .17, ~3.7 std deviations from the mean, the value of $\mathbf{H_\Delta}$ is near the margin of clear distinction, .28, ~2.4 std. deviations, from the mean of measures consistent with random walk.

As a technical note, one might question the comparison of step variances for curves with irregular point spacing. In this data set samples were collected at twice the density during the period of greatest interest, the time of rapid transition. Whether it is valid to use point spacing rather than time spacing for determining an average value of $\mathbf{H_\sigma}$ depends on whether the result is approximately a straight line since the ratios of variances would then all change in the same proportion between scales.

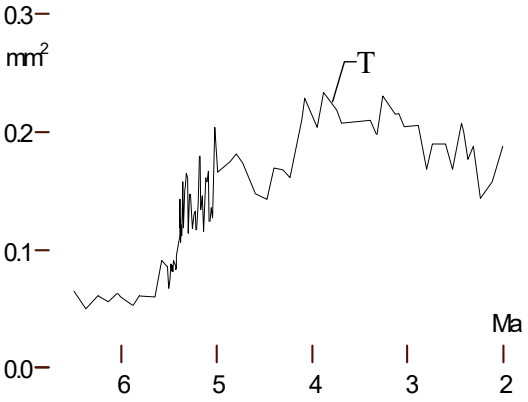


Figure 10. *Mean Profile Area of G. tumida fossils from B. Malmgren (1983) DSDP site 214*. The change in size of a common plankton over six million years recorded in 95 samples of about 50 specimens each.

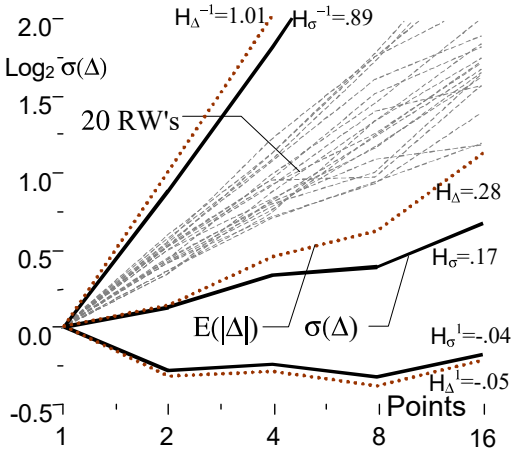


Figure 11. *Variation scaling for G. tumida size, derivative and integral, and for 20 random walks*. Comparing E(|Δ|) and σ(Δ) ratios for n=2,4,8, & 16 points. The low values of H_σ, and H_Δ indicate a presence of white noise in the data and high values of H_σ^{-1}, and H_Δ^{-1} (for integrals) indicate the noise is masking some accumulative shape.

3.4 Shape Scale Display

Having determined a predominance of symmetric noise, the G. *tumida* data may then be tentatively treated as representing a continuous curve with superimposed random noise. Figures 12 & 13 show initial noise reduction and iterated Gaussian smoothing to demonstrate the application of scale space methods to time series. The curves for figure 12 were formed by first applying $\mathbf{k}_1$, using a 17 data point spread to suppress random noise and represent the basic shape. The point spacing of the data was then regularized at .05 Myr and further shape smoothing was done with a constant 5 point (.2 Myr) Gaussian kernel.

One might question using a number of points to define the range for the initial smoothing and a fixed period for the later smoothing iterations. The initial smoothing is for the purpose of noise suppression. Once that has been done the reason for further smoothing is not noise suppression, but shape suppression, using algorithms designed for uniform point spacing. The concern is that the spacing of the data points varies widely, as more samples were taken in the period of greatest interest. The initial noise suppression based on a fixed point spread does result in non-uniform suppression of the underlying shapes that may be present. The shapes in the periods of widely spaced points (6.5 to 5.5 ma and 5 to 2 ma) are more suppressed than those in the period of most frequent points (5.5 to 5 ma). The alternate choice would have been to regularize point spacing first, resulting in exaggerated shape suppression in the period of most interest. Inspection of the results of both approaches[6] indicates that regularizing the point spacing first results in a significant misstatement of the duration of the transition event.

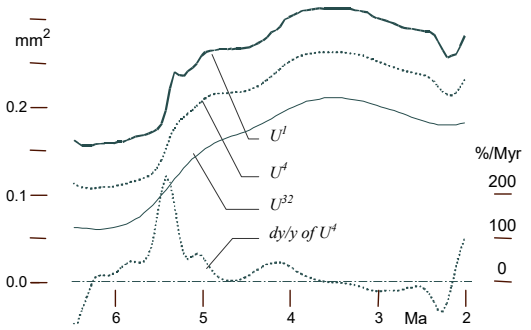


Figure 12. *Shape scales in G.tumida transition*. Thee levels of smoothing U^1, U^4, U^{32} (shown offset by .05 on the mm^2 scale for separation) along with the first log derivative of U^4. U^1 is a 17point running average of fig 9 for noise reduction. After 0.05 Myr data point spacing equalization smoothing (levels 2 to 1024) by (9 point) Gaussian kernel.

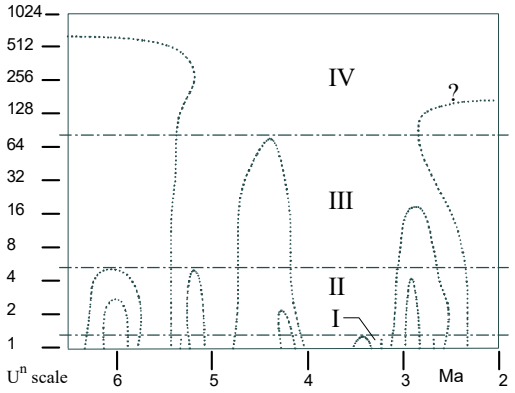


Figure 13. *Shape scale diagram*. The location of inflection points by iteration of 0.2 Myr Gaussian smoothing. General shape scale levels I, II, III, IV representing four general scales of inflection point persistence.

Whether the smoothing was actually valid depends on whether the underlying shape and it's characteristics are recognizable. This might be accomplished by matching a patterned sequence of inflection points to those in other environmental measures, demonstrating a system of corollary factors. Here, one may recognize in the derivative of U^4 in figure 12, the classic form of a triggered change of state, a clear individual event of acceleration and deceleration connecting periods with distinctly different ranges of fluctuation. The same suggestion can be gleaned from the shape scale diagram (figure 13), though this may be less clear due to the unfamiliarity of the this type of data representation.

Figure 13 shows the movement and suppression of the inflection points of the curve with iterative Gaussian smoothing. The diagram demonstrates the endurance of each inflection point and the order in which they vanish from the shape. The curves shown were drawn by plotting the location of each inflection point at each smoothing level and then manually fitting a b-spline to pass through those points. More sophisticated methods are available[7,12] but this rudimentary approach is equivalent for the present purpose. The apparent shape scale boundaries at approximately $U^{1.5}$, $U^{4.5}$, U^{64}, and U^{1024} identify groups of successively more robust inflection points.

The first two shape scales might represent real shapes in the underlying process, or residual effects of noise. It would be surprising if there were no residual effects of noise from a highly irregular data set such as this. That three pairs of inflection point traces in shape scale II drop out at about the same time suggests that they represent a true scale of secondary fluctuation in the behavior. The last inflection point to the right, starting from about 2.3 ma at U^1 is of questionable significance, being retained through 256 levels of smoothing perhaps only because the last point in the sequence was atypical. The smoothing kernel used retains the points near the ends by reducing the point range of the kernel from the normal 5 to 3 to 0, so that the end points are fixed.

As a result fluctuations near the ends are suppressed much more slowly and their retention at higher levels of smoothing less significant.

The third shape scale represents the most robust shapes in the data. The middle pair of inflection points traces back to a dip in the curve shortly after the most rapid rise, seeming to indicate that this feature is not statistical, but behavioral. Referring back to the derivative curve in figure 12, the shape scale evidence indicates that fluctuations following the singular peak in the curve at least initially reflect behavioral oscillation. At the fourth shape scale only one inflection point remains corresponding to the turning point of the apparent 'S' curve transition between relative steady states. This is the shape which makes the data interesting, positioned consistently at about 5.4 ma coinciding with the peak of the first derivative spike in U^4.

4 Discussion

The selected techniques reported on here serve to identify and interpret the invariant local structure of derivative continuity in smooth shapes. Derivative interpolation and smoothing (§2.1) is oriented to low level feature detection, regularizing the higher derivative rates of change in a sequence with minimal effect on gross shape. Gaussian shape scale analysis is oriented to hierarchical and higher level feature detection, systematically suppressing all shape in order to reveal multiple inherent scales of shape and features of shape transcending all shape scales. One of the more troublesome aspects of natural system information is the difficulty of reading behavioral states having multiple scales of irregular fluctuation. Shape scale analysis seems well suited for displaying that common characteristic of natural systems in a manner that is both directly intuitive and provides useful identifying traits for analysis.

As presented here, neither derivative reconstruction nor shape scale analysis is oriented to locating discontinuities, but to distinguishing recognizable features within smooth shape. In other hands the analysis of Gaussian shape scale has been successfully used to improve the robustness of high level corner detection in visual images[12,13] and of minutiae detection in the analysis of fingerprints[1]. To resolve low level discontinuities, say discontinuities in second derivatives as would occur at the connection between flat surfaces and connecting curves, derivative reconstruction might be augmented with a means of adaptively enhancing discontinuities. Some efforts of this kind have been noted such as the use of shock filter image processing kernels[4,14], though whether this particular method would serve the purpose indicated has not yet been determined.

Derivative reconstruction and shape scale analysis rely on one of the highly useful features of nature, that distributed physical and behavioral connections can be read from the differential structures of smooth shape. Both images and time traces, however, typically include mixed signals of various kinds, overlapping and intermittent subjects,the complex , fragmentary and indirect effects of many things. The difference scaling exponent $\mathbf{H}_{\sigma}$ (§2.2) was developed to help distinguish between data sets in which noise reduction is a valid means of generalizing shape, and those for which it is not. Only a small sample of the extensive literature on the standard scaling exponent $\mathbf{H}_{\Delta}$ and its other variations were examined, but no indication of a developed approach to interpreting mixed signals[5,11] was found.

Both indicators, however, and Hurst's original **R/S** form as well, do seem to cut through conflicting appearances to expose the same particular hidden group of natural data patterns. These include the finding by Hurst and Mandelbrot that there tends to be an essentially infinite span of statistical interdependence in natural system records[10]. It is somewhat speculative, but

this could reflect the presence of systems which have both short and long term derivative continuity, displaying multiple scales of relevant shape.

Acknowledgements

A special note of appreciation is extended to correspondents Jeff Scargle at NASA Ames Research Center for his help and suggestion of the BATSE 551 data for analysis, to Miriam Zelditch at the Univ. of Michigan Museum of Paleontology for extensive correspondence and suggestion of the G. tumida data, and Peter Roopnarine at the Department of Invertebrate Zoology & Geology, California Academy of Sciences for raising the subject of Hurst exponents in this connection. Grateful appreciation is also extended to family and friends whose long term support have made the work possible.

References

1. A. Almansa, T. Lindeberg, "Fingerprint Enhancement by Shape Adaptation of Scale-Space Operators with Automatic Scale-Selection",*Technical Report ISRN KTH/NA/R--99/01--SE* Swedish Royal Institute of Technology *(KTH) , NADA, S-100 44 Stockholm, Sweden.* submitted to *IEEE Tran. on Image Processing*, 1999 [http://www.nada.kth.se/cvap/abstracts/cvap226.html]
2. L. Alvarez, L. Mazorra, "Signal and Image Restoration Using Shock filters and Anisotropic Diffusion", *SIAM J. On Numerical Analysis*, 31 (1994) 590-605
3. F. L. Bookstein, "Random walk and the existence of evolutionary rates", *Paleobiology*, 13(4) (1987) 446-464.
4. R. Courant and H. Robbins, *What is Mathematics, An elementary Approach to Ideas and Methods*, Oxford Univ Press, 1941.
5. H. Hastings, G. Sugihara, *Fractiles, a user's guide for the natural sciences*, Oxford Science Publications, 1993
6. P. F. Henshaw, Notes & preparatory studies 1998-99
7. T. Lindeberg, *Scale Space Theory in Computer vision,* Kluwer Academic Publishers, 1994
8. T. Lindeberg, "Scale-Space Behaviour and Invariance Properties of Differential Singularities, Shape in Picture: Mathematical Description of Shape in Grey-Level Images", *NATO ASI Series F*, vol. 126, Springer-Verlag, (1994) 591-600 and [http://www.nada.kth.se/~tony /abstracts/Lin92-NATOb.html]
9. B. A. Malmgren., W. A. Berggren, and G. P. Lohmann. "Evidence for punctuated gradualism in the late neogene Globorotalia *tumida* lineage of planktonic foraminifera" *Paleobiology* V9 (1983) p377-389
10. B. Mandelbrot, "Some Long Run Properties of Geophysical Records", *Water Resoruces Research*, 5 (1969) 321-340
11. B. Mandelbrot, *Fractiles and Scaling in Finance*, Springer-Verlag, New York, 1997
12. F. Mokhtrian and A. K. Mackworth "A Theory of Multiscale Curvature-Based Shape Representation for Planar Curves", *Transactions on Pattern Analysis and Machine Intelligence*, 14-8, Aug. 1992 789-805
13. F.Mokhtrian, R. Suomela, "Robust Image Corner Detection Through Curvature Scale Space", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 20(12) (1998) 1376-1381
14. S. Osher, L. Rudin "Feature Oriented Image Enhancement using Shock Filters", *SIAM Journal on Numerical Analysys*, 27(4) (1990)

15. B. M. ter Haar Romeny, et al eds., Scale-Space '97: Proc. First Internat. Conf. on Scale-Space Theory in Computer Vision, vol. 1252 of Lecture Notes in Computer Science. Berlin: Springer Verlag, 1997; tutorial [http://www.cv.ruu.nl/Publications/pdf/vbc96tutorial.pdf]
16. P. Roopnarine, G. Byars, P Fitzgerald, "Anagenetic evolution, stratophenetic patterns, and random walk models", *Paleobiology,* 25 (1999) 41-57
17. Harry Wechsler, *Computational Vision*, Academic Press, New York, 1990
18. H. Zhao, S. Osher, B. Merriman, M. Kang "Implicit and Non-parametric Shape Reconstruction from Unorganized Data Using Variational Level Set Method", TR 98-7, UCLA Dept. of Mathematics, Los Angeles CA 90095 submitted to *Computer Vision and Image Understanding, 1999,* also [www.math.ucla.edu/applied /cam/index.html]