Reports

Synthesizing a Color Algorithm from Examples

ANYA C. HURLBERT AND TOMASO A. POGGIO*

A lightness algorithm that separates surface reflectance from illumination in a Mondrian world is synthesized automatically from a set of examples, which consist of pairs of input (intensity signal) and desired output (surface reflectance) images. The algorithm, which resembles a new lightness algorithm recently proposed by Land, is approximately equivalent to filtering the image through a center-surround receptive field in individual chromatic channels. The synthesizing technique, optimal linear estimation, requires only one assumption, that the operator that transforms input into output is linear. This assumption is true for a certain class of early vision algorithms that may therefore be synthesized in a similar way from examples. Other methods of synthesizing algorithms from examples, or "learning," such as back-propagation, do not yield a significantly better lightness algorithm.

THE PROBLEMS THAT A VISUAL SYSTEM must solve in decoding two-dimensional images into three-dimensional scenes (inverse optics problems) are difficult: the information supplied by an image is not sufficient by itself to specify a unique scene. To reduce the number of possible interpretations of images, visual systems, whether artificial or biological, must make use of natural constraints which are explicit assumptions about the physical properties of surfaces and lights. Computational vision scientists have derived effective solutions for some inverse optics problems (such as computing depth from binocular disparity) by determining the appropriate natural constraints and embedding them in algorithms. How might a visual system discover and exploit natural constraints on its own? We address a simpler question: Given only a set of examples of input images and desired output solutions, can a visual system synthesize, or "learn," the algorithm that converts input to output? We find that an algorithm for computing color in a restricted world can be constructed from examples by means of standard techniques of optimal linear estimation.

The computation of color is a prime example of the difficult problems of inverse optics. Human beings do not merely discriminate between different wavelengths of light; we assign roughly constant colors to objects even though the intensity signals they send to our eyes change as the illumination varies across space and chromatic spectrum. The computational goal underlying color constancy seems to be the extraction of invariant surface spectral reflectance properties from the image intensity signal, in which reflectance and illumination are mixed (*1*).

Lightness algorithms (*2–8*), pioneered by Land, assume that the color of an object can be specified by its lightness, or relative surface reflectance, in each of three independent chromatic channels, and that lightness is computed in the same way in each channel. Computing color is thereby reduced to extracting surface reflectance from the intensity signal in a single chromatic channel.

The image intensity signal, or more precisely, the image irradiance s' is proportional to the product of the illumination intensity e' and the surface reflectance r' in that channel:

$$s'(x,y) = r'(x,y)e'(x,y) \quad (1)$$

where x and y specify a point on the surface. This form of the intensity equation is true for a Lambertian reflectance model, in which the irradiance s' has no specular components, and for appropriately chosen color channels (*9*). Taking the logarithm of both sides converts it to a sum:

$$s(x,y) = r(x,y) + e(x,y) \quad (2)$$

where $s = \log(s')$, $r = \log(r')$, and $e = \log(e')$.

Given $s(x,y)$ alone, the problem of solving Eq. 2 for $r(x,y)$ is underconstrained. Lightness algorithms constrain the problem by restricting their domain to a world of Mondrians, two-dimensional surfaces covered with patches of random colors (*2*) (Fig. 1) and by exploiting two constraints in that world: (i) $r'(x,y)$ is uniform within patches but has sharp discontinuities at edges between patches and (ii) $e'(x,y)$ varies smoothly across the Mondrian. Under these constraints, lightness algorithms can recover a good approximation to $r(x,y)$ and so can recover lightness triplets that label roughly constant colors (*10*).

We asked whether it is possible to synthesize from examples an algorithm to extract reflectance from image irradiance, and whether the synthesized algorithm would resemble existing lightness algorithms derived from an explicit analysis of the constraints. We made one assumption, that the operator that transforms irradiance into reflectance is linear. Under that assumption, motivated by considerations discussed later, we used optimal linear estimation techniques to synthesize an operator from examples. The examples are pairs of images: an input image of a Mondrian under illumination that varies smoothly across space paired with the desired output image that displays the reflectance of the Mondrian without the illumination. The technique finds the linear estimator that best maps input into desired output, in the least-squares sense.

For computational convenience we used one-dimensional "training vectors" that represent vertical scan lines across the Mondrian images (Fig. 2). We generated many different input vectors s by adding together different random r and e vectors, according to Eq. 2. Each vector r represents a pattern of step changes across space, corresponding to one column of a reflectance image (Fig. 3, left). The step changes occur at random pixels and are of random amplitude between set minimum and maximum values. Each vector e represents a smooth gradient across space with a random offset and slope, corre-

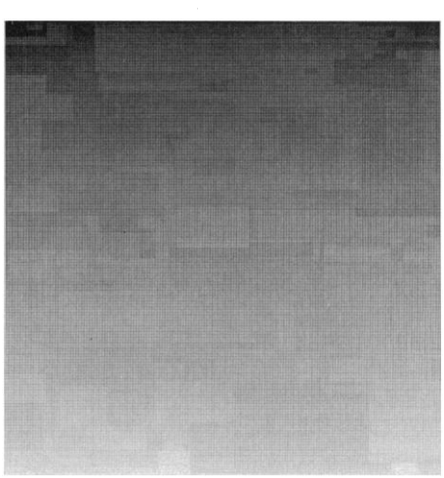

Fig. 1. A Mondrian under an illumination gradient, generated by adding together two 320 by 320 pixel images: one is the (log) reflectance image, an array of rectangles each with a different, uniform gray-level (Fig. 2A); the other is the (log) illumination image, in which the pixel values increase linearly in the same way across each column.

Artificial Intelligence Laboratory and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139.

*Also at Thinking Machines Corporation, Cambridge, MA 02142.

sponding to one column of an illumination image. We then arranged the training vectors s and r as the columns of two matrices S and R, respectively. Our goal was then to compute the optimal solution L of

$$LS = R \quad (3)$$

where L is a linear operator represented as a matrix.

It is well known that the solution of this equation that is optimal in the least-squares sense is

$$L = RS^{+} \quad (4)$$

where S^{+} is the Moore-Penrose pseudoinverse (*11*). We computed the pseudoinverse by overconstraining the problem—that is, by using many more training vectors than there are number of pixels in each vector—and by using the straightforward formula that applies in the overconstrained case: $S^{+} = S^{T}(SS^{T})^{-1}$ (*12*).

The operator L computed in this way recovers a good approximation to the correct output vector r when given a new s, not part of the training set, as input (Fig. 2C). A second operator, estimated in the same way, recovers the illumination e. Acting on a random two-dimensional Mondrian, L also yields a satisfactory approximation to the correct output image (Fig. 3, right).

Our estimation scheme has successfully synthesized an algorithm that performs the lightness computation in a Mondrian world. What is the algorithm and what is its relationship to other lightness algorithms? To answer these questions we examined the structure of the matrix L. We assumed that, although the operator is not a convolution operator, it should approximate one when applied far from the boundaries of the image. That is, in its central part, the operator should be space-invariant, performing the same action on each point in the image. Each row in the central part of L should therefore be the same as the row above but displaced by one element to the right. Inspection of the matrix confirms this expectation. To find the form of L in its center, we thus averaged the rows there, first shifting them appropriately. The result, shown in Fig. 4, is a space-invariant filter with a narrow positive peak and a broad, shallow, negative surround.

The filter our scheme synthesizes is very similar to Land's most recent retinex operator (*5*), which divides the image irradiance at each pixel by a weighted average of the irradiance at all pixels in a large surround and takes the logarithm of that result to yield lightness (*13*). The lightness triplets computed by the retinex operator agree well with human perception in a Mondrian world. The retinex operator and our matrix L both differ from Land's earlier retinex algorithms, which require a nonlinear thresholding step to eliminate smooth gradients of illumination. The shape of the filter in Fig. 4, particularly of its large surround, is also suggestive of the "nonclassical" recep-

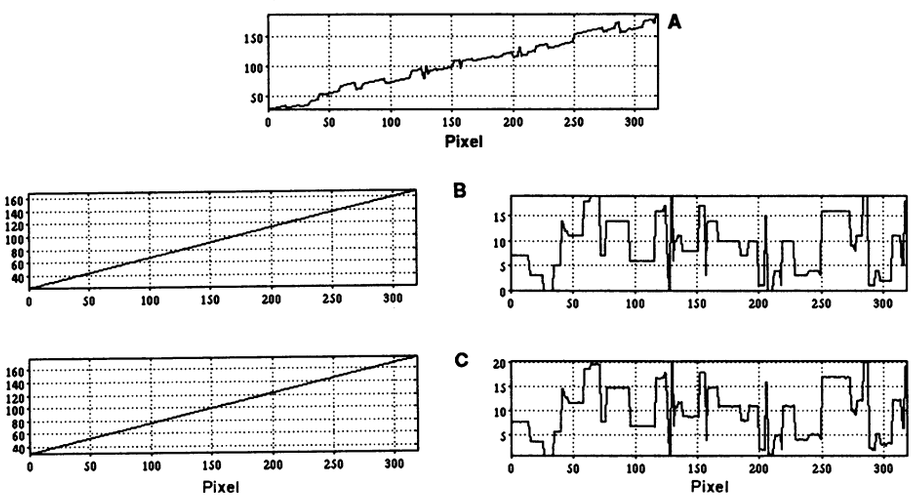

Fig. 2. (**A**) The input data, a one-dimensional vector 320 pixels long. Its random Mondrian reflectance pattern is superimposed on a linear illumination gradient with a random slope and offset. (**B**) Corresponding output solution, the illumination on the left and reflectance on the right. We used 1500 such pairs of input-output examples (each different from the others) to train the operator shown in Fig. 4. (**C**) Result obtained by the estimated operator when it acts on the input data (A), not part of the training set; illumination is on the left and the reflectance is on the right, to be compared with (B). The result is fairly typical: in some cases the prediction is even better, in others it is worse.

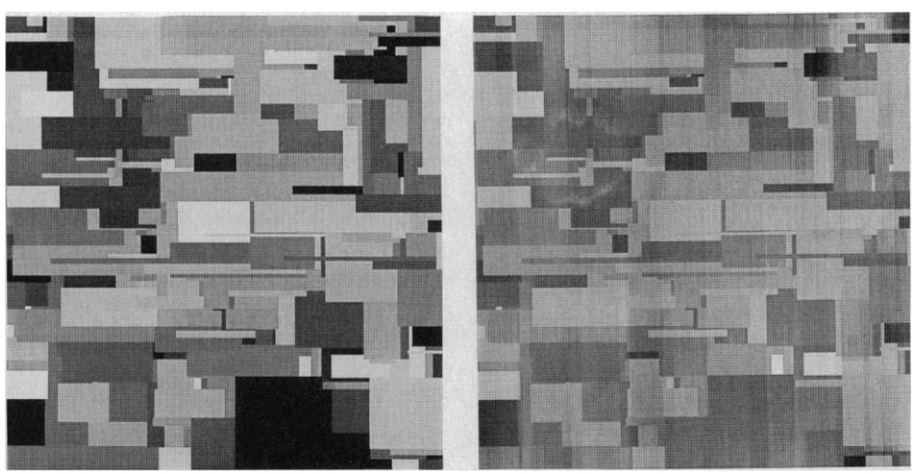

Fig. 3. (**Left**) The (log) reflectance image that is one component of the Mondrian of Fig. 1, equivalent to the Mondrian of Fig. 1 under uniform illumination. (**Right**) The (log) reflectance image that the estimated operator produces when it acts on the Mondrian of Fig. 1.

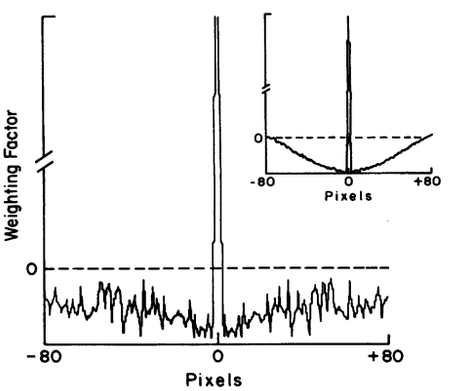

Fig. 4. The space-invariant part of the estimated operator, obtained by shifting and averaging the rows of a 160-pixel-wide central square of the matrix L, trained on a set of 1500 examples with linear illumination gradients (see Fig. 2). When logarithmic illumination gradients are used, a qualitatively similar receptive field is obtained. In a separate experiment we used a training set of one-dimensional Mondrians with either linear illumination gradients or slowly varying sinusoidal illumination components with random wavelength, phase, and amplitude. The resulting filter is shown in the inset. The surrounds of both filters extend beyond the range we can estimate reliably, the range we show here.

tive fields that have been found in V4, a cortical area implicated in mechanisms underlying color constancy (*14–17*).

The form of the space-invariant filter is similar to that derived in our earlier formal analysis of the lightness problem (*8*). It is qualitatively the same as that which results from the direct application of regularization methods exploiting the spatial constraints on reflectance and illumination described above (*9, 18, 19*). The Fourier transform of the filter of Fig. 4 is approximately a bandpass filter that cuts out low frequencies that are the result of slow gradients of illumination and preserves intermediate frequencies that are the result of step changes in reflectance. In contrast, the operator that recovers the illumination, e, takes the form of a low-pass filter. We stress that the entire operator L is not a space-invariant filter.

In this context, it is clear that the shape of the estimated operator should vary with the type of illumination gradient in the training set. We synthesized a second operator with a new set of examples that contained equal numbers of vectors with random, sinusoidally varying illumination components and vectors with random, linear illumination gradients. Whereas the first operator, synthesized from examples with strictly linear illumination gradients, has a broad negative surround that remains virtually constant throughout its extent, the new operator's surround (Fig. 4, inset) has a smaller extent and decays smoothly toward zero from its peak negative value in its center.

We also applied the operator in Fig. 4 to new input vectors in which the density and amplitude of the step changes of reflectance differed greatly from those on which the operator was trained. The operator performs well, for example, on an input vector representing one column of an image of a small patch of one reflectance against a uniform background of a different reflectance, where the entire image is under a linear illumination gradient. This result is consistent with psychophysical experiments that show that color constancy of a patch holds when its Mondrian background is replaced by an equivalent gray background (*20*).

The operator also reproduces simultaneous brightness contrast, as expected from the shape and sign of its surround. The output reflectance it computes for a patch of fixed input reflectance decreases linearly with increasing average irradiance of the input test vector in which the patch appears, in the same way that, to us, a dark patch appears darker when against a light background than against a dark one.

Our estimation procedure is motivated by our previous observation (*9, 18, 21*) that standard regularization algorithms (*19*) in early vision define linear mappings between input and output and therefore can be estimated associatively under certain conditions. The technique of optimal linear estimation that we use is closely related to optimal Bayesian estimation (*22*). If we were to assume from the start that the optimal linear operator is space-invariant, we could considerably simplify (and streamline) the computation by using standard correlation techniques (*22, 23*).

We have compared our estimation technique with other methods of "learning" a lightness algorithm. In particular, we computed the regularized pseudoinverse using gradient descent on a "neural" network (*24*) with linear units. Since the pseudoinverse is the best linear approximation in the L_2 norm, a gradient descent method that minimizes the square error between the actual output and desired output of a fully connected linear network is guaranteed to converge, albeit slowly. Thus gradient descent in weight space converged to the same result, the global minimum, as our first technique. We also compared the linear estimation technique with a "back-propagation" network: gradient descent on a two-layer network with sigmoid units (*24*) (32 inputs, 32 "hidden units," and 32 linear outputs), with training vectors 32 pixels long. The time needed for the network to converge to a stable configuration was much longer than for the linear estimator for the same set of examples. The network's performance was slightly, yet consistently, better.

We do not think that our results mean that color constancy may be learned during a critical period by biological organisms. It seems more reasonable to consider them simply as a demonstration on a toy world that in the course of evolution a visual system may recover and exploit natural constraints hidden in the physics of the world. The significance of our results lies in the facts that a simple statistical technique may be used to synthesize a lightness algorithm from examples; that the technique does as well as other techniques such as back-propagation; and that a similar technique may be used for other problems in early vision. Furthermore, the synthesized operator resembles both Land's psychophysically tested retinex operator and a neuronal nonclassical receptive field. The operator's properties suggest that simultaneous brightness (or color) contrast might be the result of the visual system's attempt to discount illumination gradients (*25*).

REFERENCES AND NOTES

1. Since we do not have perfect color constancy, our visual system must not extract reflectance exactly. The limits on color constancy might reveal limits on the underlying computation.
2. E. H. Land, *Am. Sci.* **52**, 247 (1964).
3. ——— and J. J. McCann, *J. Opt. Soc. Am.* **61**, 1 (1971).
4. E. H. Land, in *Central and Peripheral Mechanisms of Colour Vision*, T. Ottoson and S. Zeki, Eds. (Macmillan, New York, 1985), pp. 5–17.
5. ———, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3078 (1986).
6. B. K. P. Horn, *Comput. Graphics Image Proc.* **3**, 277 (1974).
7. A. Blake, in *Central and Peripheral Mechanisms of Colour Vision*, T. Ottoson and S. Zeki, Eds. (Macmillan, New York, 1985), pp. 45–59.
8. A. Hurlbert, *J. Opt. Soc. Am. A* **3**, 1684 (1986).
9. ——— and T. Poggio, *Artificial Intelligence Laboratory Memo 909* (Massachusetts Institute of Technology, Cambridge, MA, 1987).
10. The function $r'(x,y)$ can be recovered at best only to within a constant, since Eq. 1 is invariant under the transformation of r' into ar' and e' into $a^{-1}e'$, where a is any constant.
11. A. Albert, *Regression and the Moore-Penrose Pseudoinverse* (Academic Press, New York, 1972).
12. The pseudoinverse, and therefore L, may also be computed by recursive techniques that improve its form as more data become available (*11*).
13. Our synthesized filter is not exactly identical with Land's: the filter of Fig. 4 subtracts from the value at each point the average value of the logarithm of irradiance at all pixels, rather than the logarithm of the average values. The estimated operator is therefore linear in the logarithms, whereas Land's is not. The numerical difference between the outputs of the two filters is small in most cases (E. H. Land, personal communication), and both agree well with psychophysical results.
14. R. Desimone, S. J. Schein, J. Moran, L. G. Ungerleider, *Vision Res.* **25**, 441 (1985).
15. H. M. Wild, S. R. Butler, D. Carden, J. J. Kulikowski, *Nature (London)* **313**, 133 (1985).
16. S. M. Zeki, *Neurosci.* **9**, 741 (1983).
17. ———, *ibid.*, p. 767.
18. T. Poggio *et al.*, in *Proceedings of the Image Understanding Workshop*, L. Baumann, Ed. (Science Applications International Corporation, McLean, VA, 1985), pp. 25–39.
19. T. Poggio, V. Torre, C. Koch, *Nature (London)* **317**, 314 (1985).
20. A. Valberg and B. Lange-Malecki, *Invest. Ophthalmol. Visual Science Suppl.* **28**, 92 (1987).
21. T. Poggio and A. Hurlbert, *Artificial Intelligence Laboratory Working Paper 264* (Massachusetts Institute of Technology, Cambridge, MA, 1984).
22. The problem is to construct an operator L that provides the best estimation Ls of r. If the vectors s and r are sample sequences of Gaussian stochastic processes with zero mean, then the processes are fully specified by their correlation functions

$$E[ss^{\mathrm{T}}] = C_{ss}, E[rs^{\mathrm{T}}] = C_{rs} \quad (5)$$

where E indicates the expected value. The BLUE of r (*11*) is, given s,

$$r^{\mathrm{est}} = C_{rs}C_{ss}^{-1}s \quad (6)$$

which is to be compared with the regression equation

$$Ls = RS^{\mathrm{T}}(SS^{\mathrm{T}})^{-1}s \quad (7)$$

The quantities RS^{T} and SS^{T} are approximations to C_{rs} and C_{ss}, respectively, since the quantities are estimated over a finite number of observations (the training examples).
If the operator L is space invariant, the previous equation is equivalent to the formulation of an optimal linear filter also called a matched filter. Equation 7 becomes

$$R^{\mathrm{est}} = H^{\mathrm{opt}} \cdot Y \quad (8)$$

with

$$H^{\mathrm{opt}} = \Phi_{ry}/\Phi_{yy} \quad (9)$$

where Φ_{ry} is the ensemble cross power spectrum of output and inputs and Φ_{yy} is the power spectrum of the inputs. Y and R are the Fourier transforms of the inputs and estimated outputs, respectively, and H is the transfer function of the optimal filter. Both Φ_{ry}

and Φ_{yy} can be estimated from the set of r and y example pairs.
23. Estimation of the operator on two-dimensional examples is possible, but computationally very expensive if done in the same way. The present computer simulations require several hours when run on standard serial computers. The two-dimensional case will need much more time (our one-dimensional estimation scheme runs orders of magnitude faster on a CM-1 Connection Machine System with 16K-processors).
24. D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Nature (London)* **323**, 533 (1986).
25. We are grateful to E. Land, E. Hildreth, J. Little, F. Wilczek, and D. Hillis for reading the draft and for useful discussions. A. Rottenberg developed the routines for matrix operations that we used on the Connection Machine. T. Breuel wrote the back-propagation simulator.

8 September 1987; accepted 14 December 1987

Seawater Strontium Isotopes, Acid Rain, and the Cretaceous-Tertiary Boundary

J. D. MACDOUGALL

A large bolide impact at the end of the Cretaceous would have produced significant amounts of nitrogen oxides by shock heating of the atmosphere. The resulting acid precipitation would have increased continental weathering greatly and could be an explanation for the observed high ratio of strontium-87 to strontium-86 in seawater at about this time, due to the dissolution of large amounts of strontium from the continental crust. Spikes to high values in the seawater strontium isotope record at other times may reflect similar episodes.

IN RECENT YEARS MUCH PROGRESS HAS been made in two areas of research that appear unrelated: precise documentation of the strontium isotopic composition of ocean water through time and the understanding of the atmospheric geochemical effects that accompany impacts of large extraterrestrial objects on the earth. Both may be important for understanding the events that occurred at the Cretaceous-Tertiary (K-T) boundary.

Hess *et al.* (*1*) recently presented new data on the $^{87}Sr/^{86}Sr$ ratio in seawater over the past 100 million years. They analyzed selected fossil foraminifera from Deep Sea Drilling Project cores and showed that the $^{87}Sr/^{86}Sr$ ratio generally increases smoothly from near 0.7074 at 100 million years ago to the present-day seawater value of 0.7092. However, anomalously high ratios occur at the K-T boundary, forming a pronounced maximum in the otherwise relatively smooth curve. They considered whether a large impact, suggested as the cause of iridium anomalies at this boundary (*2*), could also be responsible for the strontium isotopic anomaly. Two possibilities were examined: (i) that the elevated $^{87}Sr/^{86}Sr$ ratio was due to dissolution of the bolide in seawater and (ii) that the impact ejecta (vapor and solid) were the major source of the strontium. However, Hess *et al.* (*1*) showed that although both continental crust and chondritic meteorites have high strontium isotopic ratios, these sources could not supply sufficient strontium to account for the observed seawater increase.

Prinn and Fegley (*3*) analyzed the atmospheric chemical effects of the impact of a large projectile on the earth and, following an earlier idea by Lewis *et al.* (*4*), calculated that the major effect would be the production of large amounts of nitrogen oxides (NO_x) due to shock heating of the atmosphere. As a consequence, extremely acidic precipitation would occur immediately near the impact site, and, over a more extended time period, it would occur globally. The resulting enhanced weathering of the continents would increase the supply of continental strontium to the oceans, and potentially might be a cause for the high $^{87}Sr/^{86}Sr$ ratio observed at the K-T boundary. Even under a normal weathering regime the most important factor controlling strontium isotopic variation in seawater is the isotopic composition of the riverine strontium input (*5*).

Seawater strontium isotope data from three recent papers (*1, 5, 6*) for the time period from 50 million to 80 million years ago are plotted in Fig. 1. Age assignments for most of these samples are probably accurate to 1 million to 2 million years, although there may be more uncertainty attached to some of the non–Deep Sea Drilling Project samples analyzed by DePaolo and Ingram (*6*). Agreement among the three laboratories for samples of approximately the same age is generally good, and for modern shells and seawater the bias-adjusted values agree within reported uncertainties. Thus apparently "anomalous" values (Fig. 1) are probably not due to interlaboratory effects. The high values near the K-T boundary are apparent, although there is some scatter. One sample in particular [sample P3 measured by DePaolo and Ingram (*6*) and assigned an age of 62.5 million years] that is somewhat younger than the boundary nevertheless has a high $^{87}Sr/^{86}Sr$ ratio. Comparison with the other isotopic data suggests that it may be a few million years older. The sample is a "Middle Danian" rhynchonellid brachiopod from Faxe in Denmark; this species was one of the first to appear in the Tertiary in this area (*7*). If this sample is reworked or if the age assignment is erroneously low, then the strontium isotopic spike appears to begin very abruptly at the K-T boundary, high values occur over a time period of about 2 million years, and the ratios then begin to decrease to pre-Tertiary values (Fig. 1). This is consistent with rapid introduction into the oceans of strontium

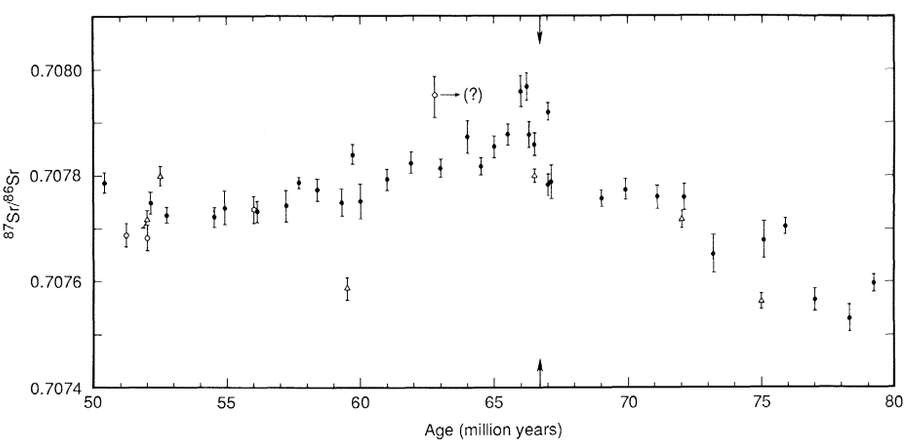

Fig. 1. Strontium isotopic composition of seawater for the period from 50 million to 80 million years ago, based on data from (*1*) (solid circles), (*5*) (open triangles), and (*6*) (open circles). Arrows on the time axis indicate the K-T boundary. The apparently anomalous sample P3 (question mark) is discussed in the text. All data are bias-adjusted to the $^{87}Sr/^{86}Sr$ value for National Bureau of Standards standard strontium given in (*1*); errors are as reported in the original works.

Scripps Institution of Oceanography, La Jolla, CA 92093.