A Global Geometric Framework for Nonlinear Dimensionality Reduction

Joshua B. Tenenbaum,[1*] Vin de Silva,[2] John C. Langford[3]

Scientists working with large volumes of high-dimensional data, such as global climate patterns, stellar spectra, or human gene distributions, regularly confront the problem of dimensionality reduction: finding meaningful low-dimensional structures hidden in their high-dimensional observations. The human brain confronts the same problem in everyday perception, extracting from its high-dimensional sensory inputs--30,000 auditory nerve fibers or 10^6 optic nerve fibers--a manageably small number of perceptually relevant features. Here we describe an approach to solving dimensionality reduction problems that uses easily measured local metric information to learn the underlying global geometry of a data set. Unlike classical techniques such as principal component analysis (PCA) and multidimensional scaling (MDS), our approach is capable of discovering the nonlinear degrees of freedom that underlie complex natural observations, such as human handwriting or images of a face under different viewing conditions. In contrast to previous algorithms for nonlinear dimensionality reduction, ours efficiently computes a globally optimal solution, and, for an important class of data manifolds, is guaranteed to converge asymptotically to the true structure.

[1] Department of Psychology and
[2] Department of Mathematics, Stanford University, Stanford, CA 94305, USA.
[3] Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15217, USA.
[*] To whom correspondence should be addressed. E-mail: jbt@psych.stanford.edu

---

A canonical problem in dimensionality reduction from the domain of visual perception is illustrated in Fig. 1A. The input consists of many images of a person's face observed under different pose and lighting conditions, in no particular order. These images can be thought of as points in a high-dimensional vector space, with each input dimension corresponding to the brightness of one pixel in the image or the firing rate of one retinal ganglion cell. Although the input dimensionality may be quite high (e.g., 4096 for these 64 pixel by 64 pixel images), the perceptually meaningful structure of these images has many fewer independent degrees of freedom. Within the 4096-dimensional input space, all of the images lie on an intrinsically three-dimensional manifold, or constraint surface, that can be parameterized by two pose variables plus an azimuthal lighting angle. Our goal is to discover, given only the unordered high-dimensional inputs, low-dimensional representations such as Fig. 1A with coordinates that capture the intrinsic degrees of freedom of a data set. This problem is of central importance not only in studies of vision (1-5), but also in speech (6, 7), motor control (8, 9), and a range of other physical and biological sciences (10-12).

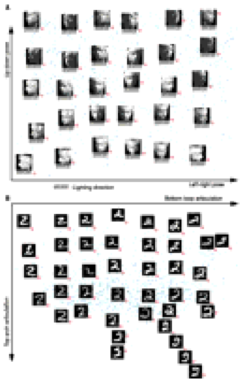

Fig. 1. (**A**) A canonical dimensionality reduction problem from visual perception. The input consists of a sequence of 4096-dimensional vectors, representing the brightness values of 64 pixel by 64 pixel images of a face rendered with different poses and lighting directions. Applied to $N = 698$ raw images, Isomap ($K = 6$) learns a three-dimensional embedding of the data's intrinsic geometric structure. A two-dimensional projection is shown, with a sample of the original input images (red circles) superimposed on all the data points (blue) and horizontal sliders (under the images) representing the third dimension. Each coordinate axis of the embedding correlates highly with one degree of freedom underlying the original data: left-right pose (*x* axis, $R = 0.99$), up-down pose (*y* axis, $R = 0.90$), and lighting direction (slider position, $R = 0.92$). The input-space distances $d_X(i,j)$ given to Isomap were Euclidean distances between the 4096-dimensional image vectors. (**B**) Isomap applied to $N = 1000$ handwritten "2"s from the MNIST database (40). The two most significant dimensions in the Isomap embedding, shown here, articulate the major features of the "2": bottom loop (*x* axis) and top arch (*y* axis). Input-space distances $d_X(i,j)$ were measured by tangent distance, a metric designed to capture the invariances relevant in handwriting recognition (41). Here we used ϵ-Isomap (with $\epsilon = 4.2$) because we did not expect a constant dimensionality to hold over the whole data set; consistent with this, Isomap finds several tendrils projecting from the higher dimensional mass of data and representing successive exaggerations of an extra stroke or ornament in the digit.
[View Larger Version of this Image (39K GIF file)]

The classical techniques for dimensionality reduction, PCA and MDS, are simple to implement, efficiently computable, and guaranteed to discover the true structure of data lying on or near a linear subspace of the high-dimensional input space (13). PCA finds a low-dimensional embedding of the data points that best preserves their variance as measured in the high-dimensional input space. Classical MDS finds an embedding that preserves the interpoint distances, equivalent to PCA when those distances are Euclidean. However, many data sets contain essential nonlinear structures that are invisible to PCA and MDS (4, 5, 11, 14). For example, both methods fail to detect the true degrees of freedom of the face data set (Fig. 1A), or even its intrinsic three-dimensionality (Fig. 2A).

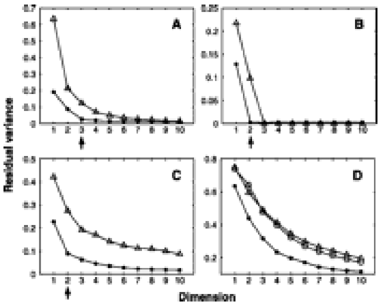

Fig. 2. The residual variance of PCA (open triangles), MDS [open triangles in (A) through (C); open circles in (D)], and Isomap (filled circles) on four data sets (42). (**A**) Face images varying in pose and illumination (Fig. 1A). (**B**) Swiss roll data (Fig. 3). (**C**) Hand images varying in finger extension and wrist rotation (20). (**D**) Handwritten "2"s (Fig. 1B). In all cases, residual variance decreases as the dimensionality *d* is increased. The intrinsic dimensionality of the data can be estimated by looking for the "elbow" at which this curve ceases to decrease significantly with added dimensions. Arrows mark the true or approximate dimensionality, when known. Note the tendency of PCA and MDS to overestimate the dimensionality, in contrast to Isomap.
[View Larger Version of this Image (23K GIF file)]

Here we describe an approach that combines the major algorithmic features of PCA and MDS--computational efficiency, global optimality, and asymptotic convergence guarantees--with the flexibility to learn a broad class of nonlinear manifolds. Figure 3A illustrates the challenge of nonlinearity with data lying on a two-dimensional "Swiss roll": points far apart on the underlying manifold, as measured by their geodesic, or shortest path, distances, may appear deceptively close in the high-dimensional input space, as measured by their straight-line Euclidean distance. Only the geodesic distances reflect the true low-dimensional geometry of the manifold, but PCA and MDS effectively see just the Euclidean structure; thus, they fail to detect the intrinsic two-dimensionality (Fig. 2B).

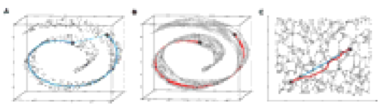

Fig. 3. The "Swiss roll" data set, illustrating how Isomap exploits geodesic paths for nonlinear dimensionality reduction. (**A**) For two arbitrary points (circled) on a nonlinear manifold, their Euclidean distance in the high-dimensional input space (length of dashed line) may not accurately reflect their intrinsic similarity, as measured by geodesic distance along the low-dimensional manifold (length of solid curve). (**B**) The neighborhood graph *G* constructed in step one of Isomap (with $K = 7$ and $N = 1000$ data points) allows an approximation (red segments) to the true geodesic path to be computed efficiently in step two, as the shortest path in *G*. (**C**) The two-dimensional embedding recovered by Isomap in step three, which best preserves the shortest path distances in the neighborhood graph (overlaid). Straight lines in the embedding (blue) now represent simpler and cleaner approximations to the true geodesic paths than do the corresponding graph paths (red). [View Larger Version of this Image (23K GIF file)]

Our approach builds on classical MDS but seeks to preserve the intrinsic geometry of the data, as captured in the geodesic manifold distances between all pairs of data points. The crux is estimating the geodesic distance between faraway points, given only input-space distances. For neighboring points, input-space distance provides a good approximation to geodesic distance. For faraway points, geodesic distance can be approximated by adding up a sequence of "short hops" between neighboring points. These approximations are computed efficiently by finding shortest paths in a graph with edges connecting neighboring data points.

The complete isometric feature mapping, or Isomap, algorithm has three steps, which are detailed in Table 1. The first step determines which points are neighbors on the manifold M, based on the distances $d_X(i,j)$ between pairs of points i,j in the input space X. Two simple methods are to connect each point to all points within some fixed radius ϵ, or to all of its K nearest neighbors (15). These neighborhood relations are represented as a weighted graph G over the data points, with edges of weight $d_X(i,j)$ between neighboring points (Fig. 3B).

Table 1. The Isomap algorithm takes as input the distances $d_X(i,j)$ between all pairs i,j from N data points in the high-dimensional input space X, measured either in the standard Euclidean metric (as in Fig. 1A) or in some domain-specific metric (as in Fig. 1B). The algorithm outputs coordinate vectors $\mathbf{y}_i$ in a d-dimensional Euclidean space Y that (according to Eq. 1) best represent the intrinsic geometry of the data. The only free parameter (ϵ or K) appears in Step 1.

Step		
1	Construct neighborhood graph	Define the graph G over all data points by connecting points i and j if [as measured by $d_X(i,j)$] they are closer than ϵ (ϵ-Isomap), or if i is one of the K nearest neighbors of j (K-Isomap). Set edge lengths equal to $d_X(i,j)$.
2	Compute shortest paths	Initialize $d_G(i,j) = d_X(i,j)$ if i,j are linked by an edge; $d_G(i,j) = \infty$ otherwise. Then for each value of $k = 1, 2, ..., N$ in turn, replace all entries $d_G(i,j)$ by $\min\{d_G(i,j), d_G(i,k) + d_G(k,j)\}$. The matrix of final values $D_G = \{d_G(i,j)\}$ will contain the shortest path distances between all pairs of points in G (16, 19).
3	Construct d-dimensional embedding	Let λ_p be the p-th eigenvalue (in decreasing order) of the matrix $\tau(D_G)$ (17), and v_p^i be the i-th component of the p-th eigenvector. Then set the p-th component of the d-dimensional coordinate vector $\mathbf{y}_i$ equal to $\sqrt{\lambda_p}\, v_p^i$.

In its second step, Isomap estimates the geodesic distances $d_M(i,j)$ between all pairs of points on the manifold M by computing their shortest path distances $d_G(i,j)$ in the graph G. One simple algorithm (16) for finding shortest paths is given in Table 1.

The final step applies classical MDS to the matrix of graph distances $D_G = \{d_G(i,j)\}$, constructing an embedding of the data in a d-dimensional Euclidean space Y that best preserves the manifold's estimated intrinsic geometry (Fig. 3C). The coordinate vectors $\mathbf{y}_i$ for points in Y are chosen to minimize the cost function

$$E = \| \tau(D_G) - \tau(D_Y) \|_{L^2} \tag{1}$$

where D_Y denotes the matrix of Euclidean distances $\{d_Y(i,j) = \|\mathbf{y}_i - \mathbf{y}_j\|\}$ and $\|A\|_{L^2}$ the L^2 matrix norm $\sqrt{\Sigma_{i,j} A_{ij}^2}$. The τ operator converts distances to inner products (17), which uniquely characterize the geometry of the data in a form that supports efficient optimization. The global minimum of Eq. 1 is achieved by setting the coordinates $\mathbf{y}_i$ to the top d eigenvectors of the matrix $\tau(D_G)$ (13).

As with PCA or MDS, the true dimensionality of the data can be estimated from the decrease in error as the dimensionality of Y is increased. For the Swiss roll, where classical methods fail, the residual variance of Isomap correctly bottoms out at $d = 2$ (Fig. 2B).

Just as PCA and MDS are guaranteed, given sufficient data, to recover the true structure of linear manifolds, Isomap is guaranteed asymptotically to recover the true dimensionality and geometric structure of a strictly larger class of nonlinear manifolds. Like the Swiss roll, these are manifolds whose intrinsic geometry is that of a convex region of Euclidean space, but whose ambient geometry in the high-dimensional input space may be highly folded, twisted, or curved. For non-Euclidean manifolds, such as a hemisphere or the surface of a doughnut, Isomap still produces a globally optimal low-dimensional Euclidean representation, as measured by Eq. 1.

These guarantees of asymptotic convergence rest on a proof that as the number of data points increases, the graph distances $d_G(i,j)$ provide increasingly better approximations to the intrinsic geodesic distances $d_M(i,j)$, becoming arbitrarily accurate in the limit of infinite data (18, 19). How quickly $d_G(i,j)$ converges to $d_M(i,j)$ depends on certain parameters of the manifold as it lies within the high-dimensional space (radius of curvature and branch separation) and on the density of points. To the extent that a data set presents extreme values of these parameters or deviates from a uniform density, asymptotic convergence still holds in general, but the sample size required to estimate geodesic distance accurately may be impractically large.

Isomap's global coordinates provide a simple way to analyze and manipulate high-dimensional observations in terms of their intrinsic nonlinear degrees of freedom. For a set of synthetic face images, known to have three degrees of freedom, Isomap correctly detects the dimensionality (Fig. 2A) and separates out the true underlying factors (Fig. 1A). The algorithm also recovers the known low-dimensional structure of a set of noisy real images, generated by a human hand varying in finger extension and wrist rotation (Fig. 2C) (20). Given a more complex data set of handwritten digits, which does not have a clear manifold geometry, Isomap still finds globally meaningful coordinates (Fig. 1B) and nonlinear structure that PCA or MDS do not detect (Fig. 2D). For all three data sets, the natural appearance of linear interpolations between distant points in the low-dimensional coordinate space confirms that Isomap has captured the data's perceptually relevant structure (Fig. 4).

Fig. 4. Interpolations along straight lines in the Isomap coordinate space

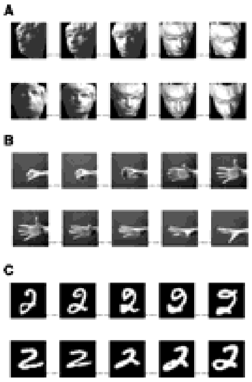

Fig. 4. Interpolations along straight lines in the Isomap coordinate space (analogous to the blue line in Fig. 3C) implement perceptually natural but highly nonlinear "morphs" of the corresponding high-dimensional observations (43) by transforming them approximately along geodesic paths (analogous to the solid curve in Fig. 3A). (**A**) Interpolations in a three-dimensional embedding of face images (Fig. 1A). (**B**) Interpolations in a four-dimensional embedding of hand images (20) appear as natural hand movements when viewed in quick succession, even though no such motions occurred in the observed data. (**C**) Interpolations in a six-dimensional embedding of handwritten "2"s (Fig. 1B) preserve continuity not only in the visual features of loop and arch articulation, but also in the implied pen trajectories, which are the true degrees of freedom underlying those appearances.

[View Larger Version of this Image (46K GIF file)]

Previous attempts to extend PCA and MDS to nonlinear data sets fall into two broad classes, each of which suffers from limitations overcome by our approach. Local linear techniques (21-23) are not designed to represent the global structure of a data set within a single coordinate system, as we do in Fig. 1. Nonlinear techniques based on greedy optimization procedures (24-30) attempt to discover global structure, but lack the crucial algorithmic features that Isomap inherits from PCA and MDS: a noniterative, polynomial time procedure with a guarantee of global optimality; for intrinsically Euclidean manifolds, a guarantee of asymptotic convergence to the true structure; and the ability to discover manifolds of arbitrary dimensionality, rather than requiring a fixed d initialized from the beginning or computational resources that increase exponentially in d.

Here we have demonstrated Isomap's performance on data sets chosen for their visually compelling structures, but the technique may be applied wherever nonlinear geometry complicates the use of PCA or MDS. Isomap complements, and may be combined with, linear extensions of PCA based on higher order statistics, such as independent component analysis (31, 32). It may also lead to a better understanding of how the brain comes to represent the dynamic appearance of objects, where psychophysical studies of apparent motion (33, 34) suggest a central role for geodesic transformations on nonlinear manifolds (35) much like those studied here.

REFERENCES AND NOTES

1. M. P. Young and S. Yamane, *Science* **256**, 1327 (1992) [ISI][Medline].
2. R. N. Shepard, *Science* **210**, 390 (1980) [ISI].
3. M. Turk and A. Pentland, *J. Cogn. Neurosci.* **3**, 71 (1991) [ISI].
4. H. Murase and S. K. Nayar, *Int. J. Comp. Vision* **14**, 5 (1995) [ISI].
5. J. W. McClurkin, L. M. Optican, B. J. Richmond, T. J. Gawne, *Science* **253**, 675 (1991) [ISI][Medline].
6. J. L. Elman and D. Zipser, *J. Acoust. Soc. Am.* **83**, 1615 (1988) [ISI][Medline].
7. W. Klein, R. Plomp, L. C. W. Pols, *J. Acoust. Soc. Am.* **48**, 999 (1970) [ISI][Medline].
8. E. Bizzi, F. A. Mussa-Ivaldi, S. Giszter, *Science* **253**, 287 (1991) [ISI][Medline].
9. T. D. Sanger, *Adv. Neural Info. Proc. Syst.* **7**, 1023 (1995) .

10. J. W. Hurrell, *Science* **269**, 676 (1995) [ISI].
11. C. A. L. Bailer-Jones, M. Irwin, T. von Hippel, *Mon. Not. R. Astron. Soc.* **298**, 361 (1997) [CrossRef][ISI].
12. P. Menozzi, A. Piazza, L. Cavalli-Sforza, *Science* **201**, 786 (1978) [ISI][Medline].
13. K. V. Mardia, J. T. Kent, J. M. Bibby, *Multivariate Analysis*, (Academic Press, London, 1979).
14. A. H. Monahan, *J. Clim.*, in press.
15. The scale-invariant K parameter is typically easier to set than ϵ, but may yield misleading results when the local dimensionality varies across the data set. When available, additional constraints such as the temporal ordering of observations may also help to determine neighbors. In earlier work (36) we explored a more complex method (37), which required an order of magnitude more data and did not support the theoretical performance guarantees we provide here for ϵ- and K-Isomap.
16. This procedure, known as Floyd's algorithm, requires $O(N^3)$ operations. More efficient algorithms exploiting the sparse structure of the neighborhood graph can be found in (38).
17. The operator τ is defined by $\tau(D) = -HSH/2$, where S is the matrix of squared distances $\{S_{ij} = D_{ij}^2\}$, and H is the "centering matrix" $\{H_{ij} = \delta_{ij} - 1/N\}$ (13).
18. Our proof works by showing that for a sufficiently high density (α) of data points, we can always choose a neighborhood size (ϵ or K) large enough that the graph will (with high probability) have a path not much longer than the true geodesic, but small enough to prevent edges that "short circuit" the true geometry of the manifold. More precisely, given arbitrarily small values of λ_1, λ_2, and μ, we can guarantee that with probability at least $1 - \mu$, estimates of the form

$$(1-\lambda_1)d_M(i,j) \le d_G(i,j) \le (1+\lambda_2)d_M(i,j)$$

will hold uniformly over all pairs of data points i,j. For ϵ-Isomap, we require

$$\epsilon \le (2/\pi)r_0\sqrt{24\lambda_1},\ \ \epsilon < s_0,$$

$$\alpha > [\log(V/\mu\eta_d(\lambda_2\epsilon/16)^d)]/\eta_d(\lambda_2\epsilon/8)^d$$

where r_0 is the minimal radius of curvature of the manifold M as embedded in the input space X, s_0 is the minimal branch separation of M in X, V is the (d-dimensional) volume of M, and (ignoring boundary effects) η_d is the volume of the unit ball in Euclidean d-space. For K-Isomap, we let ϵ be as above and fix the ratio $(K + 1)/\alpha = \eta_d(\epsilon/2)^d/2$. We then require

$$e^{-(K+1)/4} \le \mu\eta_d(\epsilon/4)^d/4V,$$

$$(e/4)^{(K+1)/2} \le \mu\eta_d(\epsilon/8)^d/16V,$$

$$\alpha > [4\log(8V/\mu\eta_d(\lambda_2\epsilon/32\pi)^d)]/\eta_d(\lambda_2\epsilon/16\pi)^d$$

The exact content of these conditions--but not their general form--depends on the particular technical assumptions we adopt. For details and extensions to nonuniform densities, intrinsic

curvature, and boundary effects, see http://isomap.stanford.edu.
19. In practice, for finite data sets, $d_G(i,j)$ may fail to approximate $d_M(i,j)$ for a small fraction of points that are disconnected from the giant component of the neighborhood graph *G*. These outliers are easily detected as having infinite graph distances from the majority of other points and can be deleted from further analysis.
20. The Isomap embedding of the hand images is available at *Science* Online at www.sciencemag.org/cgi/content/full/290/5500/2319/DC1. For additional material and computer code, see http://isomap.stanford.edu.
21. R. Basri, D. Roth, D. Jacobs, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (1998), pp. 414-420.
22. C. Bregler and S. M. Omohundro, *Adv. Neural Info. Proc. Syst.* **7**, 973 (1995) .
23. G. E. Hinton, M. Revow, P. Dayan, *Adv. Neural Info. Proc. Syst.* **7**, 1015 (1995) .
24. R. Durbin and D. Willshaw, *Nature* **326**, 689 (1987) [ISI][Medline].
25. T. Kohonen, *Self-Organisation and Associative Memory* (Springer-Verlag, Berlin, ed. 2, 1988), pp. 119-157.
26. T. Hastie and W. Stuetzle, *J. Am. Stat. Assoc.* **84**, 502 (1989) [ISI].
27. M. A. Kramer, *AIChE J.* **37**, 233 (1991) [ISI].
28. D. DeMers and G. Cottrell, *Adv. Neural Info. Proc. Syst.* **5**, 580 (1993) .
29. R. Hecht-Nielsen, *Science* **269**, 1860 (1995) [ISI].
30. C. M. Bishop, M. Svensén, C. K. I. Williams, *Neural Comp.* **10**, 215 (1998) [Abstract].
31. P. Comon, *Signal Proc.* **36**, 287 (1994) [ISI].
32. A. J. Bell and T. J. Sejnowski, *Neural Comp.* **7**, 1129 (1995) [Abstract].
33. R. N. Shepard and S. A. Judd, *Science* **191**, 952 (1976) [ISI][Medline].
34. M. Shiffrar and J. J. Freyd, *Psychol. Science* **1**, 257 (1990) [ISI].
35. R. N. Shepard, *Psychon. Bull. Rev.* **1**, 2 (1994) [ISI].
36. J. B. Tenenbaum, *Adv. Neural Info. Proc. Syst.* **10**, 682 (1998) .
37. T. Martinetz and K. Schulten, *Neural Netw.* **7**, 507 (1994) [ISI].
38. V. Kumar, A. Grama, A. Gupta, G. Karypis, *Introduction to Parallel Computing*: *Design and Analysis of Algorithms* (Benjamin/Cummings, Redwood City, CA, 1994), pp. 257-297.
39. D. Beymer and T. Poggio, *Science* **272**, 1905 (1996) [Abstract].
40. Available at www.research.att.com/~yann/ocr/mnist
41. P. Y. Simard, Y. LeCun, J. Denker, *Adv. Neural Info. Proc. Syst.* **5**, 50 (1993) .
42. In order to evaluate the fits of PCA, MDS, and Isomap on comparable grounds, we use the residual variance 1 - $R^2(\hat{D}_M, D_Y)$. D_Y is the matrix of Euclidean distances in the low-dimensional embedding recovered by each algorithm. $\hat{D}_M$ is each algorithm's best estimate of the intrinsic manifold distances: for Isomap, this is the graph distance matrix D_G; for PCA and MDS, it is the Euclidean input-space distance matrix D_X (except with the handwritten "2"s, where MDS uses the tangent distance). *R* is the standard linear correlation coefficient, taken over all entries of $\hat{D}_M$ and D_Y.
43. In each sequence shown, the three intermediate images are those closest to the points 1/4, 1/2, and 3/4 of the way between the given endpoints. We can also synthesize an explicit mapping from input space *X* to the low-dimensional embedding *Y*, or vice versa, using the coordinates of corresponding points $\{\mathbf{x}_i, \mathbf{y}_i\}$ in both spaces provided by Isomap together with standard supervised learning techniques (39).
44. Supported by the Mitsubishi Electric Research Laboratories, the Schlumberger Foundation, the

NSF (DBS-9021648), and the DARPA Human ID program. We thank Y. LeCun for making available the MNIST database and S. Roweis and L. Saul for sharing related unpublished work. For many helpful discussions, we thank G. Carlsson, H. Farid, W. Freeman, T. Griffiths, R. Lehrer, S. Mahajan, D. Reich, W. Richards, J. M. Tenenbaum, Y. Weiss, and especially M. Bernstein.

10 August 2000; accepted 21 November 2000

This article has been cited by other articles:

- Agrafiotis, D. K., Xu, H. (2002). A self-organizing principle for learning nonlinear manifolds. *Proc. Natl. Acad. Sci. U. S. A.* 99: 15869-15872 [Abstract] [Full Text]
- Mott, R., Schultz, J., Bork, P., Ponting, C. P. (2002). Predicting Protein Cellular Localization Using a Domain Projection Method. *Genome Res.* 12: 1168-1174 [Abstract] [Full Text]
- Levin, D. N. (2002). Sensor-independent stimulus representations. *Proc. Natl. Acad. Sci. U. S. A.* 99: 7346-7351 [Abstract] [Full Text]
- Bolhuis, P. G., Chandler, D., Dellago, C., Geissler, P. L. (2002). TRANSITION PATH SAMPLING: Throwing Ropes Over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* 53: 291-318 [Abstract] [Full Text]
- Sotiriou, C., Khanna, C., Jazaeri, A. A., Petersen, D., Liu, E. T. (2002). Core Biopsies Can Be

Used to Distinguish Differences in Expression Profiling by cDNA Microarrays. *J Mol Diagn* 4: 30-36 [Abstract] [Full Text]
- Koehl, P., Levitt, M. (2002). Improved recognition of native-like protein structures using a family of designed sequences. *Proc. Natl. Acad. Sci. U. S. A.* 99: 691-696 [Abstract] [Full Text]
- Balasubramanian, M., Schwartz, E. L., Tenenbaum, J. B., de Silva, V., Langford, J. C. (2002). The Isomap Algorithm and Topological Stability. *Science* 295: 7a-7 [Full Text]

Related articles in Science:

COGNITION:
The Manifold Ways of Perception
H. Sebastian Seung and Daniel D. Lee
Science 2000 290: 2268-2269. (in Perspectives) [Summary] [Full Text]

Nonlinear Dimensionality Reduction by Locally Linear Embedding
Sam T. Roweis and Lawrence K. Saul
Science 2000 290: 2323-2326. (in Reports) [Abstract] [Full Text]

Volume 290, Number 5500, Issue of 22 Dec 2000, pp. 2319-2323.

PAGE TOP