Neurocomputing 74 (2011) 941–948

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A new quality assessment criterion for nonlinear dimensionality reduction

Deyu Meng [a], Yee Leung [b,*], Zongben Xu [a]

[a] Institute for Information and System Sciences, Xi'an Jiaotong University, Xi'an 710049, PR China
[b] Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong, PR China

ARTICLE INFO

Article history:
Received 23 June 2010
Received in revised form
9 October 2010
Accepted 26 October 2010
Communicated by C. Fyfe
Available online 16 December 2010

Keywords:
Manifold learning
Nonlinear dimensionality reduction
Pattern recognition
Quality assessment

ABSTRACT

A new quality assessment criterion for evaluating the performance of the nonlinear dimensionality reduction (NLDR) methods is proposed in this paper. Differing from the current quality assessment criteria focusing on the local-neighborhood-preserving performance of the NLDR methods, the proposed criterion capitalizes on a new aspect, the global-structure-holding performance, of the NLDR methods. By taking both properties into consideration, the intrinsic capability of the NLDR methods can be more faithfully reflected, and hence more rational measurement for the proper selection of NLDR methods in real-life applications can be offered. The theoretical argument is supported by experiment results implemented on a series of benchmark data sets.

1. Introduction

Data collected for practical applications in various fields, such as biological sciences and multimedia information processing, are often of very high dimensionality that causes great difficulties in data mining and knowledge discovery. However, the high-dimensional data are often sampled as a probability distribution on a smooth manifold with intrinsic low dimensionality. Hence, if one can find the essential low-dimensional representations (or embeddings) of the high-dimensional raw data, it will greatly facilitate further processing and analysis, such as pattern recognition, visualization and query, of such data. And, this becomes the main issue of nonlinear dimensionality reduction (NLDR).

Various methods for NLDR have been developed over the years. They include isometric feature mapping or Isomap [1], locally linear embedding or LLE [2], Laplacian eigenmap [3], local tangent space alignment or LTSA [4], Hessian LLE [5], maximum variance unfolding or MVU [6], locally linear coordination or LLC [7], neighborhood preserving embedding or NPE [8], linearity preserving projection or LPP [9], Sammon mapping, generative topographic mapping, the visualization induced SOM [10,11], and others [12,13]. These methods attempt to capture the low-dimensional representations of the original high-dimensional data, so that the representations can optimally preserve the local configurations of the nearest neighbors of the raw data, and can further recover the global geometry underlying the original data manifold.

The aforementioned methods all attempt to accomplish similar task and have their own advantages and disadvantages due to different characteristics of their construction principles. How to design a reasonable strategy to select an appropriate NLDR method for a given high-dimensional data set, and how to establish a rational quality assessment criterion for appropriately evaluating the performance of different NLDR methods, have thus become important issues in data-mining research [14–16].

The current state-of-the-art on this issue can be mainly represented by the following approaches. The first approach evaluates the performance of a NLDR method by looking at the value of the objective function it optimizes after convergence [3,17]. Although the approach allows us to quantitatively assess the qualities of multiple implementations (e.g., with different parameter values) of similar methods, it makes the comparison of different methods unfair. The second approach takes the reconstruction error of a NLDR method as its quality assessment criterion [18]. Obviously, this criterion is reasonable and universal. Yet it requires the availability of the bidirectional mappings between the high-dimensional raw data space and low-dimensional representational space in a closed form. However, most NLDR methods are nonparametric, and they only provide values of the mapping from the raw data space to the representational space for merely the known parameters. This leads to the infeasibility of this criterion in many practical NLDR cases. The third approach utilizes an indirect performance index, such as classification accuracy, to evaluate the quality of the representations calculated by a NLDR method [19]. Yet the approach only works for data with special properties, such as the labeled data. The well-known Spearmans rho measurement has also been employed to estimate how well the corresponding low-dimensional projection preserves the order of the pairwise distances of the high-dimensional data points [20]. For high-dimensional data

* Corresponding author.
E-mail address: yeeleung@cuhk.edu.hk (Y. Leung).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.10.011

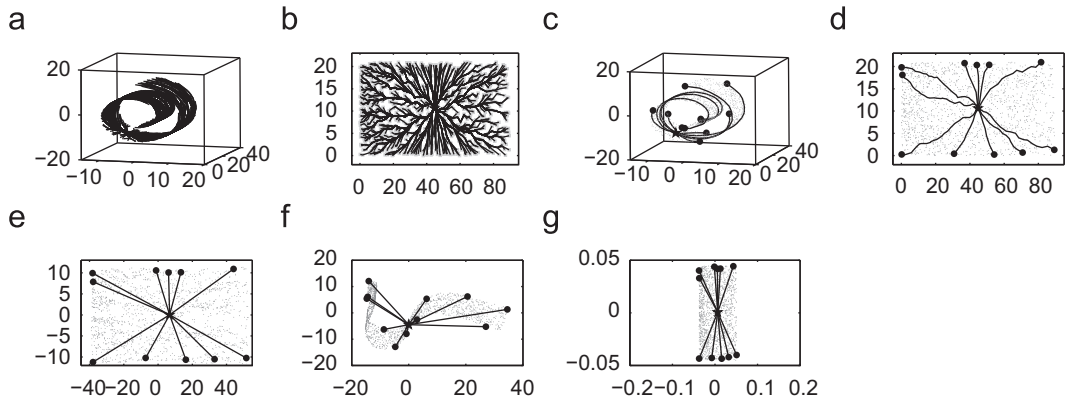

Fig. 1. (a) The shortest path tree (SPT) of the Swiss-roll data. The star point is its root vertex, calculated by Formula (4). (c) The main branches of the SPT. The solid dots denote the leafs of these branches. (b) and (d) The precise 2-D unwrapped data projections of (a) and (c). (e)–(g) The 2-D embeddings of the Swiss-roll data obtained from Isomap, MVU, and LTSA, respectively. The solid lines denote the corresponding main branches of these embeddings. It should be noted that the aspect ratios of (d), (e), (g) are all set as 4:1 for easy visualization of the similar and dissimilar transforming scales along the *x* and *y* axes of the embeddings of (e) and (g) compared to the original manifold figure (d), respectively.

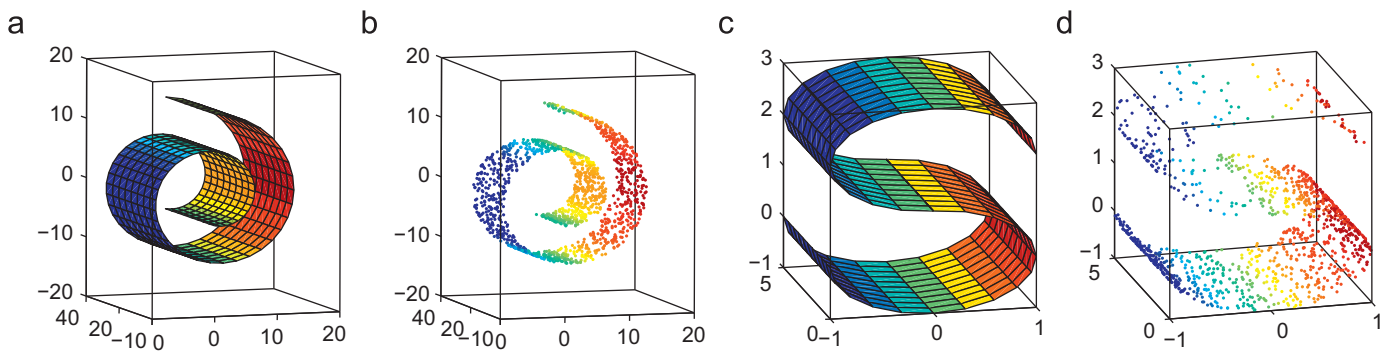

Fig. 2. (a) Swiss-roll manifold, wrapped by a 2-D 4:1 rectangle. (b) Data set A with 1500 points generated from the manifold (a). (c) S-curve manifold, wrapped by a 2-D 2:1 rectangle. (d) Data set B with 1200 points generated from the manifold (b). The upper and lower parts of the data set consist of 200 and 1000 points, respectively.

lying on a low-dimensional manifold, however, such a quality assessment generally cannot take effect, since only small pairwise distances between nearby data can be preserved in the projection while large distances between faraway ones incline to highly deviate from their real geodesic distance along the underlying manifold.

The more recent approach to this problem sticks to the intrinsic goal of NLDR and attempts to faithfully assess how well the original data structure is preserved by the representations obtained from a NLDR method. The quality criterion is generally of wider applicability, and are simpler and more rational in its construction principle than the above three approaches [21]. The latest developments along this line of reasoning include the local continuity meta-criterion or LCMC [14,22], the trustworthiness and continuity measure or T&C [16,23], the mean relative rank error or MRRE [15,21,24], etc. All of these criteria focus on the preservation of the local geometric structures of the data, and analyze the situation in the neighborhood pre-specified around each of the entire data [1–7] in the NLDR implementation. Accordingly, these criteria are always able to successfully assess the local-neighborhood-preserving performance of the employed NLDR methods.

Nevertheless, criteria employed in the current studies are still not satisfactory because they neglect the global-structure-holding capability of the utilized NLDR methods. That is, the representational set can have a perfect local-neighborhood-preserving property but it does not mean that it is of good global-geometry-holding capability (both constitute the goal of NLDR method). On one hand, even for a perfect local-neighborhood-preserving embedding set, it may encounter unexpected distortions in different parts of its global shape. This can easily be observed in Fig. 1(f), which shows the embeddings of the Swiss-roll data (as depicted in Fig. 2(b)), with intrinsic 2-D 4:1 rectangle figure, calculated by the MVU method.[1] On the other hand, some spectral methods tend to yield low-dimensional embedding set with different transforming scales in its various feature coordinates, as clearly shown in Fig. 1(g), which depicts the embeddings of the Swiss-roll data by utilizing the LTSA method for NLDR implementation. For applications such as pattern recognition, the aforementioned irregularities, i.e., large differences in global shape or aspect ratios, pose serious problems on any classifier because the global metric relationship is greatly changed. Hence, it is more reasonable to evaluate the capability of a NLDR method by considering its global-structure-holding performance as well as its local-neighborhood-preserving performance.

The motivation of this paper is to propose a new approach to more properly assess the performance of the NLDR methods. In particular, a new quality assessment criterion is constructed for evaluating the global-geometry-holding property of the embeddings calculated by the employed NLDR techniques. Then by simultaneously considering both their local-neighborhood-preserving and the global-geometry-holding properties, the performance of the employed NLDR techniques can be more faithfully and comprehensively assessed. The rationality of the proposed

[1] The good local-neighborhood-preserving performance of a NLDR method on a given data set can be visualized by the continuously changing colors of the local areas of the embeddings calculated by this method, and can also be quantitatively reflected by the corresponding larger values of LCMC, T&C, and MRRE measurements.

approach is substantiated by numerical simulations performed on a series of benchmark data sets.

The general idea and the implementation details of the new quality assessment criterion are first introduced in Section 2. The effectiveness of the proposed criterion is then verified in Section 3 by applying it to evaluate the performance of various dimensionality reduction methods on a series of benchmark data sets. The paper is concluded by a summary and outlook for future research.

2. Local and global quality assessments for NLDR

As aforementioned, current quality assessments mainly focus on the evaluation of the local-neighborhood-preserving performance of the utilized NLDR method. The purpose of this study is to construct a new measure to further assess the global-structure-holding capability of the employed method, so that by integrating it with the local measure to give more comprehensive evaluation for the performance of the NLDR method. In what follows, we first give a brief review of three recent local assessment methods, and then propose the global assessment method and the algorithm for its calculation.

2.1. Local quality assessment for NLDR

We first review in this section three of the recently presented quality assessment criteria for NLDR, namely the LCMC [14,22], the T&C [16,23], and the MRRE [15,21,24].

Denote the original data set as $X=\{x_i\}_{i=1}^{l}$, and the corresponding representational set calculated by the employed NLDR method as $Y=\{y_i\}_{i=1}^{l}$. Then the LCMC is defined as

$$Q_L = 1-\frac{1}{lk}\sum_{i=1}^{l}\left(|\aleph_k^X(i)\bigcap\aleph_k^Y(i)|-\frac{k^2}{l-1}\right), \tag{1}$$

where k is the pre-specified neighborhood size, $\aleph_k^X(i)$ is the index set of x_i's k-NN, and $\aleph_k^Y(i)$ is the index set of y_i's k-NN. Here k-NN represents the k nearest neighbors of a datum. By computing the average overlapping between two k-NN neighboring sets of the original and the representational sets, the LCMC criterion gives a general measure for the local faithfulness of the calculated embeddings.

The T&C measure involves two evaluations, the trustworthiness measure and the continuity measure, defined, respectively, as

$$M_T = 1-\frac{2}{lk(2l-3k-1)}\sum_{i=1}^{l}\sum_{j\in U_k(i)\setminus V_k(i)}(r(i,j)-k),$$

$$M_C = 1-\frac{2}{lk(2l-3k-1)}\sum_{i=1}^{l}\sum_{j\in V_k(i)\setminus U_k(i)}(\hat{r}(i,j)-k),$$

where k is the neighborhood size, $r(i,j)$ ($\hat{r}(i,j)$) is the rank of x_j (y_j) in the ordering according to the distance from $x_i(y_i)$ in the original (representational) space, and $U_k(i)$ ($V_k(i)$) is the set of those data samples that are in k-NN of $x_i(y_i)$ in the representational (original) space. In particular, M_T measures the trustworthiness degree that data points originally farther away enter the neighborhood of a sample in the embeddings; as a comparison, M_C evaluates the continuity degree that data points that are originally in the neighborhood are pushed farther away in data representations. The T&C measure is then defined as

$$Q_T = \alpha M_T + (1-\alpha)M_C, \tag{2}$$

where $\alpha\in[0,1]$ is the compromise parameter. It should be noted that in T&C measure, the trade-off between these two terms, tunable by a parameter α, governs the trade-off between trustworthiness and continuity. By properly pre-specifying α, the measure so evaluated can reflect the consistency between the local neighborhoods of the original data and the corresponding ones of the embeddings calculated by the utilized NLDR method.

The MRRE relies on the principle very similar to that of the T&C, also including two elements defined as

$$W_T = 1-\frac{1}{H_k}\sum_{i=1}^{N}\sum_{j\in U_k(i)}\frac{|r(i,j)-\hat{r}(i,j)|}{r(i,j)},$$

$$W_C = 1-\frac{1}{H_k}\sum_{i=1}^{N}\sum_{j\in V_k(i)}\frac{|r(i,j)-\hat{r}(i,j)|}{\hat{r}(i,j)},$$

where k is the neighborhood size, and $H_k = Nl\sum_{i=1}^{k}(l-2i+1)/i$ is the normalizing factor. The MRRE is of the form

$$Q_M = \beta W_T + (1-\beta)W_C, \tag{3}$$

where $\beta\in[0,1]$ is the compromise parameter. The intrinsic difference between the MRRE and the T&C is that the former considers all the k-NN samples in the representational (original) space, while the latter focuses on the k-NN of the samples in the representational (original) space but not in the original (representational) space. Also their different weights on the trustworthy and continuity components always bring more contrastive while instable performance of the MRRE measure than that of the T&C measure, as can be observed in the experiments of Section 3.

By utilizing the above criteria, the local-neighborhood-preserving performance of the employed NLDR methods can be efficiently assessed. An effective method for evaluating the global-structure-holding performance of NLDR is to be explained in the next section.

2.2. Global quality assessment for NLDR

The quality of a NLDR method can be assessed by how well the calculated embeddings can hold the global structure of the original data. Specifically, it can be evaluated by the transforming scales of the embedding figure along various directions compared to the original data manifold. If the embeddings well preserve the original global manifold figure, the transforming scales should be similar along all directions of the embeddings. On the contrary, if the embedding set encounters distortion of the global shape or aspect ratio of the original data manifold, it tends to be of distinct transforming scales along different orientations compared to the original manifold, as clearly shown in Fig. 1(f) and (g). Accordingly, by quantitatively evaluating the difference of the transforming scales of the embedding set compared with the original data manifold along various directions, the global-structure-holding quality of the utilized NLDR method can then be explicitly assessed. To this aim, two problems need to be settled: (i) how to acquire the global structure information of the original and the embedding data along different corresponding directions? and (ii) how to calculate the in-between transforming scales and how to quantify the difference of these scales?

To facilitate our discussion on the strategy of attaining the global structure information of the original data set along various orientations, we first introduce some relevant knowledge on shortest path tree (SPT) utilized in our analysis. A SPT is a rooted spanning tree of a connected graph, whose branches are taken as the shortest paths from the root vertex to all other vertices of the graph. Each vertex of a SPT is connected to a single vertex immediately following it and leads in turn to the root vertex on the shortest path. Obviously, the SPT is designated as long as the root vertex is selected. Without any prior knowledge about the underlying manifold, it is preferable to select the data center as the root. An easy way to obtain the approximate center of the data is to utilize their approximate circumcenter. For a bounded closed set

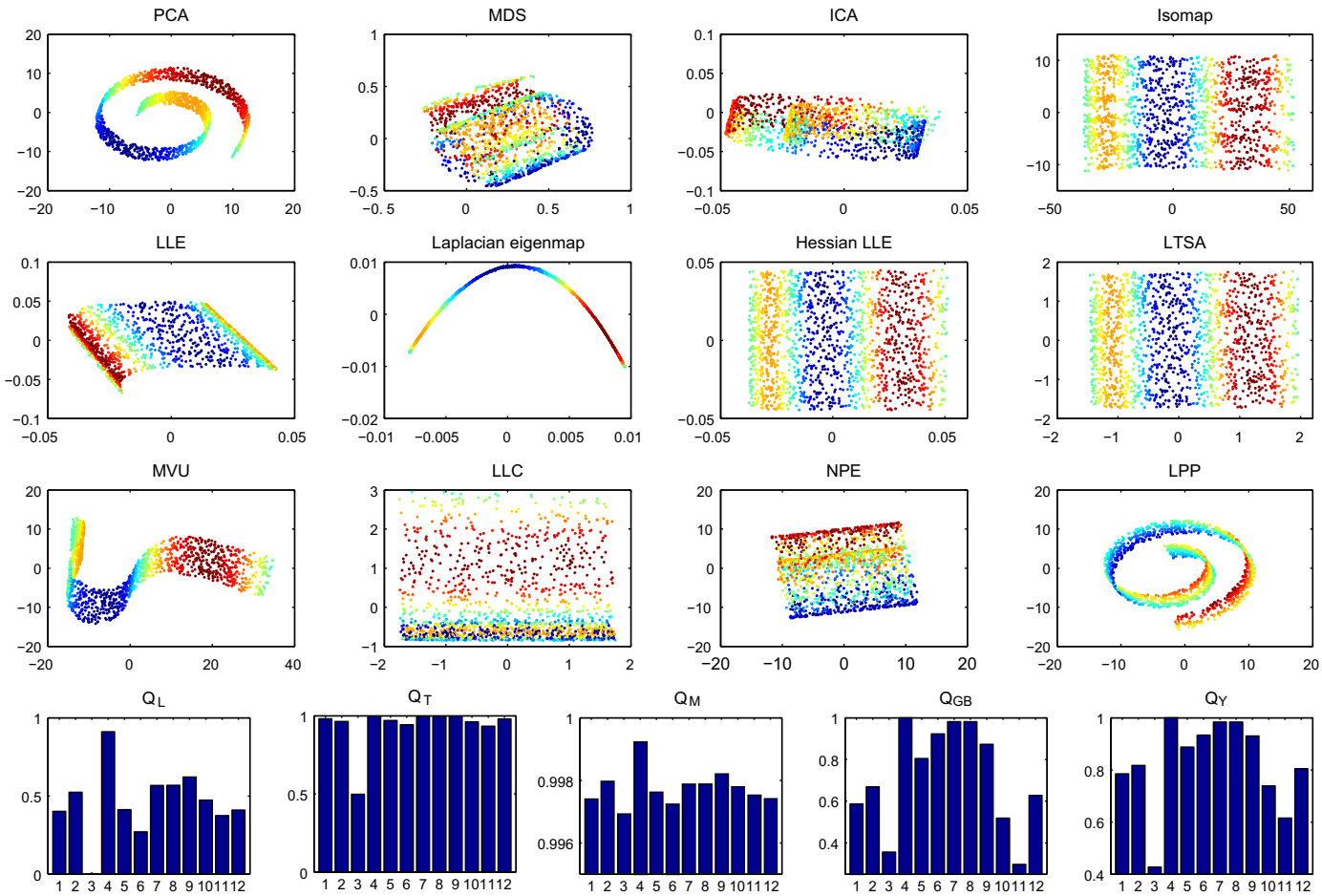

Fig. 3. 2-D embeddings of Data set A calculated by (1) PCA, (2) MDS, (3) ICA, (4) Isomap, (5) LLE, (6) Laplacian eigenmap, (7) Hessian LLE, (8) LTSA, (9) MVU, (10) LLC, (11) NPE, (12) LPP, respectively. Figures in the last row depict the corresponding quality assessments Q_L, Q_T, Q_M, Q_{GB}, Q_Y for the performance of these methods, respectively.

Ω, its circumcenter is defined as

$$Center = \underset{x \in \Omega}{\operatorname{argmin}}\left(\max_{y \in \Omega}(\|x-y\|)\right).$$

Correspondingly, the approximate circumcenter of the data set X can be obtained as

$$x_c = \underset{x_j \in X}{\operatorname{argmin}}\left(\max_{x_k \in X}(D(j,k))\right), \tag{4}$$

where $D(j,k)$ is the specified distance measure between data points x_j and x_k. Note that the optimization problem (4) can be solved by any sorting algorithms, such as the well-known heap selection algorithm [25]. Utilizing x_c so calculated as the root vertex, the SPT of the data set can then be designated.

According to the above knowledge, the global structure information of the original data set X along various orientations can be evaluated through the following three steps: First, generate the k-NN neighborhood graph $G=(V,E)$, where the vertex set V consists of the given data set X, and the edge set E contains the k-NN edges of all vertices; second, construct the SPT superimposed on X from the neighborhood graph G, as depicted in Fig. 1(a) (in the original space) and (b) (in the projection space)[2]; and third, yield the main branches of the SPT so constructed, as shown in Fig. 1(c) (in the original space) and (d) (in the projection space). Specifically, since there are k neighbors around the estimated center (the root vertex) x_c of X, and each branch of the SPT definitely passes through one neighbor of the root x_c, all branches are naturally clustered into k categories by taking the neighbors they pass by as cluster labels. Then it is reasonable to select the longest branch from each category as the representation of that category. Evidently, the branch so selected approximately depicts the orientation tendency of all branches in this category directed from the root to their leaf vertices. Subsequently, the k branches so constructed are configured as the main branches of the SPT, and the lengths (denoted as $DX=\{dx_i\}_{i=1}^k$) of these main branches, i.e., the shortest paths between the root and the corresponding leafs (denoted as $LX=\{lx_i\}_{i=1}^k$), naturally reflect the global structure of the data manifold along different directions.

Based on the evaluation of the global structure information of X, the corresponding structure of the embedding set Y can be correspondingly assessed. Denote the points of Y corresponding to x_c and $LX=\{lx_i\}_{i=1}^k$ as y_c and $LY=\{ly_i\}_{i=1}^k$, respectively. Corresponding to the length sequence $DX=\{dx_i\}_{i=1}^k$ between x_c and $LX=\{lx_i\}_{i=1}^k$, denote the sequence $DY=\{dy_i\}_{i=1}^k$ as the distances between y_c and $LY=\{ly_i\}_{i=1}^k$. Similar to DX, the sequence DY also correspondingly represents the global shapes of Y along different directions, as clearly depicted in Fig. 1(e)–(g).

Our goal is then to evaluate the global-structure-holding degree of the embedding data Y by calculating the difference of the transforming scales between X and Y along various directions, i.e., between the sequences DX and DY. This aim can approximately be attained by

[2] The root vertex of the SPT is specified as the approximate center calculated by Formula (4). The distance $D(i,j)$ in the formula is employed as the length of the shortest path, i.e., the approximate geodesic distance, between x_i and x_j.

computing the dissimilarity between the ranking orders of DX and DY, which is just the purpose for which the well-known Spearman's rank-order correlation coefficient is designed [26]. Specifically, the coefficient is computed as follows: The sequences DX and DY are first converted into their rankings, denoted as $R_{DX}=\{r_i\}_{i=1}^{k}$ and $R_{DY}=\{\tilde{r}_i\}_{i=1}^{k}$. Then Spearman's coefficient is given by

$$Q_{GB}=1-\frac{6\sum_{i=1}^{k}d_i^2}{N_k}, \tag{5}$$

where $d_i=r_i-\tilde{r}_i$, and $N_k=2k(k^2-1)$ is the normalization factor. The coefficient so calculated then rationally assesses the quality of the global-structure-holding property of the embedding set Y, i.e., the global-structure-holding performance of the employed NLDR method.

Based on the aforementioned, the algorithm for evaluating the global-structure-holding performance of the NLDR method can then be proposed as follows:

Algorithm for global quality assessment of NLDR

Input: Original data set $X=\{x_i\}_{i=1}^{l}\subset R^n$; embedding data set $Y=\{y_i\}_{i=1}^{l}\subset R^d$ $(d<n)$ calculated by some NLDR method; neighborhood size k.

Step I. Construct the k-NN neighborhood graph of X, and compute the approximate center x_c of X according to Formula (4), and then generate the SPT of the graph by taking x_c as the root vertex.

Step II. Generate the main branches of the SPT by the method introduced in Section 2.2. Denote the leaf vertices of these branches as $LX=\{lx_i\}_{i=1}^{k}$, and their lengths as $DX=\{dx_i\}_{i=1}^{k}$.

Step III. Find the corresponding points of x_c and $LX=\{lx_i\}_{i=1}^{k}$ in Y, denoted as y_c and $LY=\{ly_i\}_{i=1}^{k}$, respectively. Let the sequence $DY=\{dy_i\}_{i=1}^{k}$ be the distances between y_c and $LY_k=\{ly_i\}_{i=1}^{k}$.

Step IV. Calculate Spearman's coefficient Q_{GB} of the sequences DX and DY by utilizing Formula (5).

Output: The final quality assessment of the global-structure-holding performance of the utilized NLDR method: Q_{GB}.

The proposed criterion, Q_{GB}, presents a new aspect (the global-structure-holding degree), which is not considered by current quality assessment methods, to measure the performance of the NLDR method. In practice, different criteria measure different aspects of a NLDR method and more information is contained in examining both the local and global qualities. Thus both local and global quality assessments should be involved to more comprehensively evaluate the performance of the NLDR method. In our experiments, we employ a simple and practical way to get a single evaluation value for easy quality assessment of the utilized NLDR techniques, i.e.,

$$Q_Y=\mu Q_{GB}+(1-\mu)Q_{LC}, \tag{6}$$

where Q_{LC} is any local quality assessment obtained by the approaches introduced in Section 2.1, and $\mu\in[0,1]$ is a tuning parameter to balance the local measure Q_{LC} and the global one Q_{GB} in quality assessment.

It should be noted that to calculate the proposed quality assessment criterion, an appropriate neighborhood size k needs

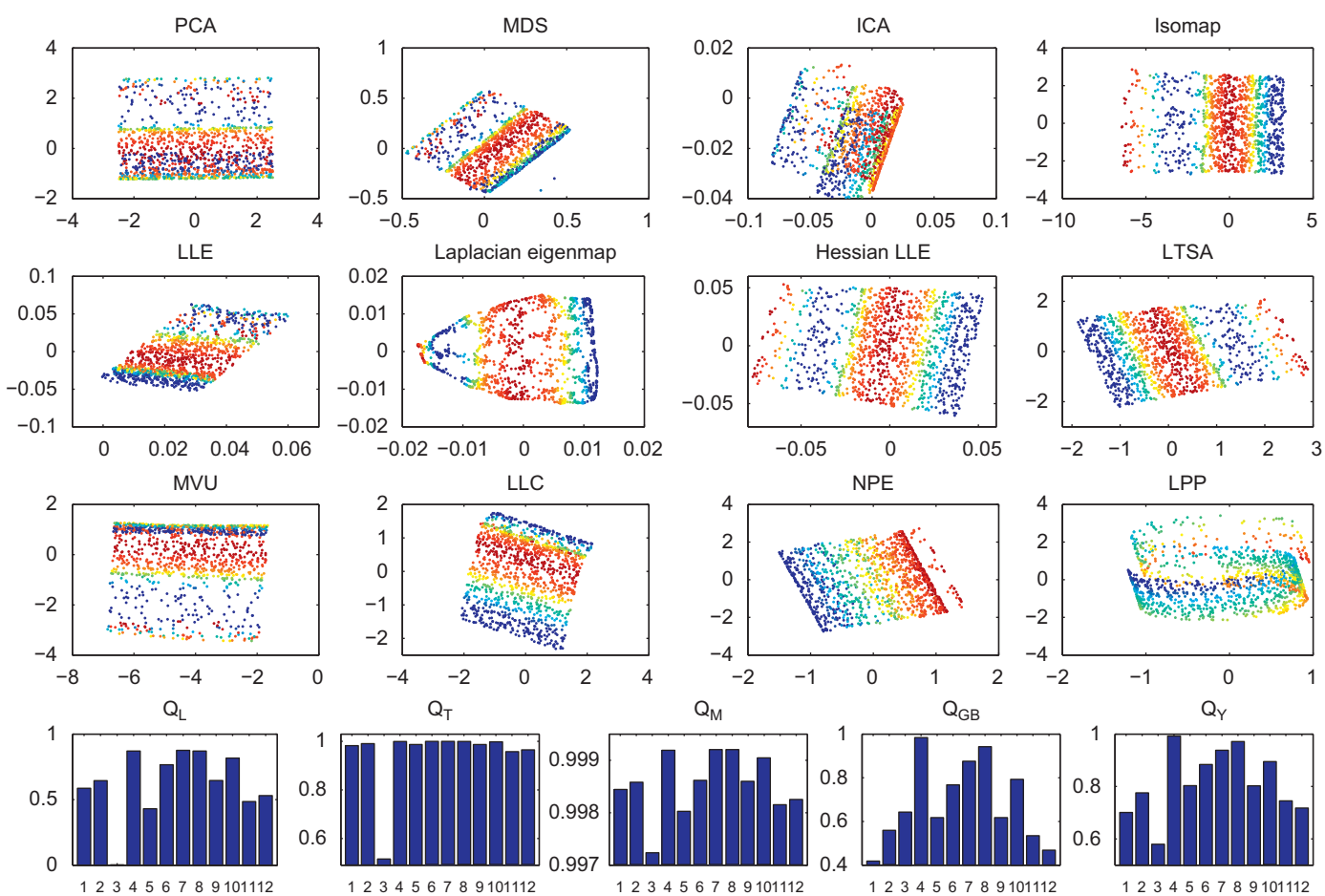

Fig. 4. 2-D embeddings of Data B calculated by the 12 dimensionality reduction methods, respectively. Figures in the last row depict the corresponding quality assessments Q_L, Q_T, Q_M, Q_{GB}, Q_Y for the performance of these methods, respectively.

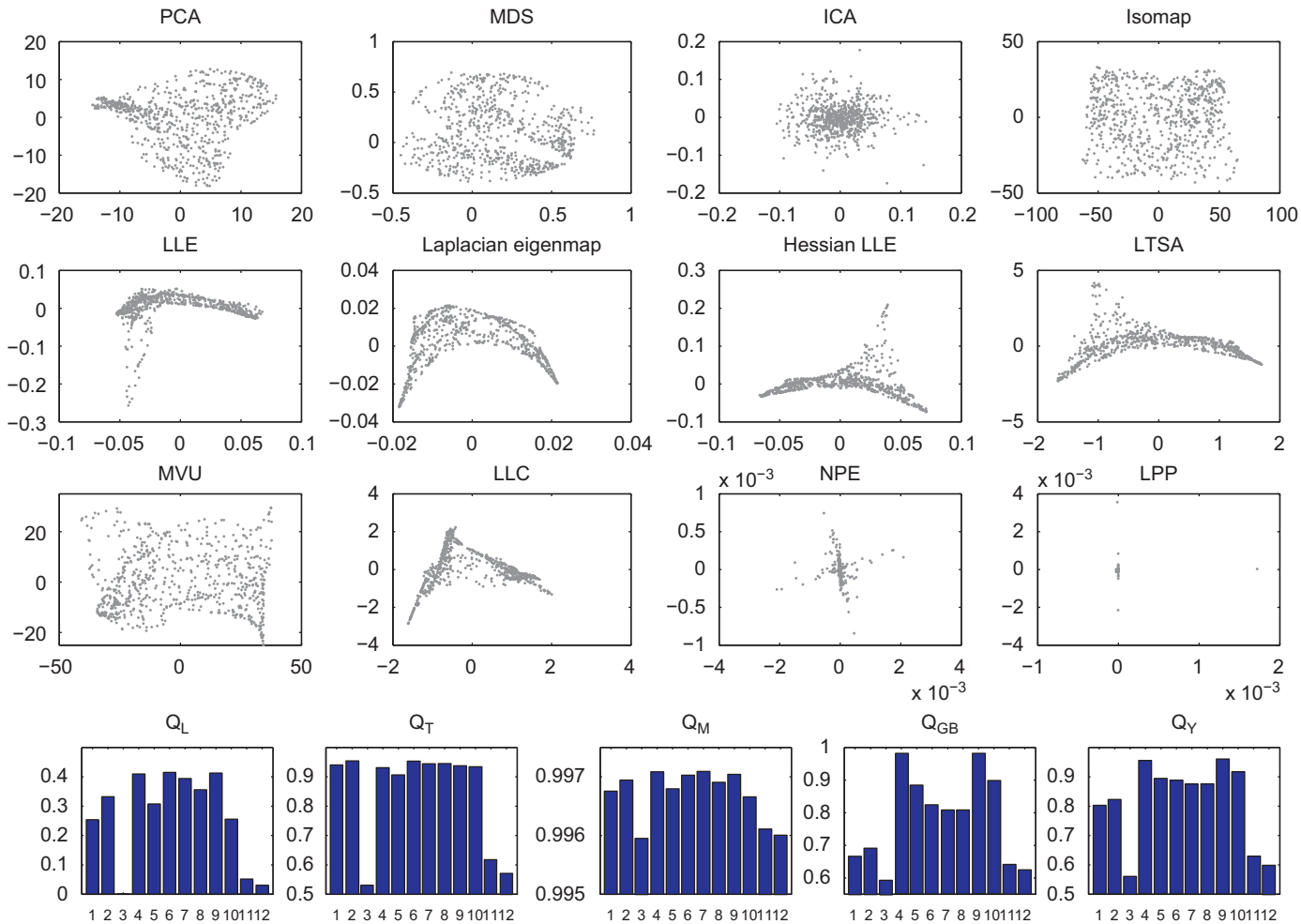

Fig. 5. 2-D embeddings of Data C calculated by the 12 dimensionality reduction methods, respectively. Figures in the last row depict the corresponding quality assessments Q_L, Q_T, Q_M, Q_{GB}, Q_Y for the performance of these methods, respectively.

to be pre-specified. A number of methods have been developed to select a proper neighborhood size for NLDR implementation. For instance, a "trial-and-error" method based on the trade-off between two cost functions has been proposed in [18]. The disadvantage of the method is that it needs to repeatedly implement the embedding algorithm, making it a little costly [27]. In [28], a neighborhood contraction and expansion method was also constructed to adaptively and efficiently calculate the neighborhood sizes at each point, and their average value can then be specified as a reasonable neighborhood size in a global aspect. This method is hence adopted in our algorithm for selecting the neighborhood size k.

Using a series of benchmark data sets in NLDR research as test examples, the effectiveness of the proposed quality assessment criterion is to be evaluated in the next section.

3. Experiments

Two synthetic and one image data sets were employed to evaluate the performance of the proposed quality assessment criterion. The synthetic Data set A and Data set B, as depicted in Fig. 2(b) and (d) with 1500 and 1200 points, respectively, were generated from the classical Swiss-roll and S-curve manifolds, as shown in Fig. 2(a) and (c), respectively. It should be noted that Data set B is of different densities in its upper and lower parts, consisting of 200 and 1000 points, respectively. The image data set (Data set C) contains 698 4096-dimensional vectors, representing the brightness values of 64 by 64 pixel images of a face with different poses and lighting directions. All of these data sets are the most commonly used benchmark data in NLDR research [1–7,13] and were thus taken as the empirical basis of our simulations. Furthermore, 12 dimensionality reduction methods, including three linear techniques: PCA [29], MDS [30], ICA [31], and nine nonlinear methods: Isomap, LLE, Laplacian eigenmap, LTSA, Hessian LLE, MVU, LLC, NPE, and LPP, were applied to these data sets in our simulations. For each 2-D embedding set of the three benchmark data sets calculated by the 12 methods, the proposed global quality assessment Q_{GB}, as computed in Formula (5), was calculated, and the local quality assessments, including Q_L, Q_T, Q_M as computed in Formulas (1), (2), (3), respectively, were also made. The results are depicted in Figs. 3, 4 and 5, respectively. Especially, the integrated quality assessment criterion Q_Y (Formula (6)), by taking Q_T as the Q_{LC} criterion (Formula (2)) and setting the value of the compromise parameter as 1/2, i.e., let Q_Y as the average of the local and global quality measurements,[3] were also demonstrated. The embeddings of the data sets are also depicted in the corresponding figures to give the overall visualization of the performance of the involved methods.

Compared to the local quality assessments Q_L, Q_T, and Q_M, it can be observed in these figures that by integrating the global measurement Q_{GB}, Q_Y more properly evaluates the performance of the dimensionality

[3] It should be noted that the presetting of μ (and furthermore, α, β) remains to be a subjective matter, and its specification still depends on the user perspective to emphasize on the local or global (and the trustworthiness or continuity) quality of a method.

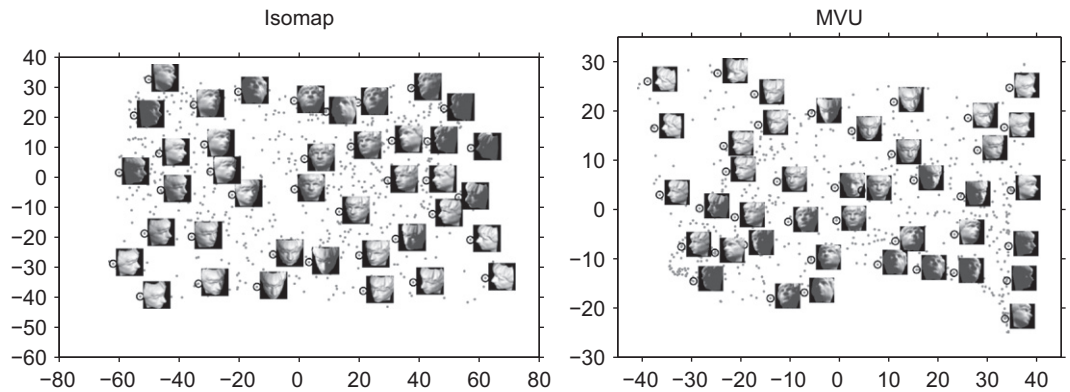

Fig. 6. 2-D embeddings of Data C calculated by Isomap and MVU. The representative images are shown next to the circled points in different parts of the space.

reduction methods. For instance, each of the three local quality assessments of the embeddings of Data set A calculated by MVU attains a larger value than that obtained by Hessian LLE and LTSA. However, it is seen that the embedding set obtained by MVU encounters a significant distortion problem in its global configuration. While in comparison, those calculated by Hessian LLE and LTSA finely recover the intrinsic rectangle figure, and faithfully keep the local neighborhood structures, of the original data manifold. The latter two methods thus possess better dimensionality reduction qualities than the former one. This is consistent with the quality assessment results of the criterion Q_Y. Besides, based on the three local quality assessments of the embeddings of Data set B, the performance of Isomap is not as good as that of Hessian LLE and LTSA. However, compared to these two methods which encounter aspect ratio abnormities on their global embedding figure, the embeddings calculated by Isomap more faithfully preserve the intrinsic 2:1 rectangle shape of the original data. As a whole, Isomap performs better than Hessian LLE and LTSA, agreeing with the qualities evaluated by the Q_Y calculated on these methods. Furthermore, for the NLDR methods applied to Data set C, both Isomap and MVU are not the most preferable methods according to the qualities assessed by Q_L, Q_T, and Q_M. Yet by visualization, it can be observed that Isomap and MVU are actually the best performers of all 12 methods, similar to what is evaluated by the criterion Q_Y. The reason is threefold: Firstly, the embedding data points are uniformly distributed and well organized in a global aspect. Secondly, nearby points in the embeddings correspond to similar images in a local aspect, as can be observed in Fig. 6. Thirdly, each coordinate axis of the embeddings highly accords with one degree representational freedom underlying the original data. Specifically, the x and y axes represent the degrees of the up-down and left-right facing poses of the face images, respectively, as depicted in Fig. 6. All these results demonstrate the more faithful quality assessment of Q_Y, which integratively considers both the proposed global measure Q_{GB} and the current local one, on evaluating the performance of the utilized NLDR techniques.

4. Conclusion

A new quality assessment criterion for evaluating the performance of the NLDR method has been proposed in this paper. Different from the conventional quality assessments that focus on the local-neighborhood-preserving performance of the NLDR method, the proposed assessment method further considers its global-structure-holding capability. Since both properties intrinsically reflect the capability of a NLDR method, the criterion by combining both global and local quality assessments generally attains more accurate quality assessment than the conventional ones. The effectiveness of the proposed criterion has been experimentally supported by its outstanding performance on a series of benchmark data sets.

Further investigations in our future research include evaluating the qualities of the performance of different NLDR methods by estimating their quality assessment tendency under a certain range of the k values, and substantiating the effectiveness of the proposed criterion by more practical applications.

Acknowledgments

The authors are very grateful to the anonymous reviewers for their valuable comments and suggestions that help us to clarify and improve various aspects of the paper. This research was supported by the National Grand Fundamental Research 973 Program of China under Grant no. 2007CB311002 and the China NSFC project under Contract 60905003.

References

[1] J.B. Tenenbaum, V. de Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (5500) (2000) 2319–2323.
[2] S.T. Roweis, L.K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (5500) (2000) 2323–2326.
[3] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, Neural Computation 15 (6) (2003) 1373–1396.
[4] Z.Y. Zhang, H.Y. Zha, Principal manifolds and nonlinear dimension reduction via local tangent space alignment, Technical Report CSE-02-019, CSE, Penn State University, 2002.
[5] D.L. Donoho, C. Grimes, Hessian eigenmaps: locally linear embedding techniques for high-dimensional data, Proceedings of the National Academy of Sciences 100 (10) (2003) 5591–5596.
[6] K.Q. Weinberger, L.K. Saul, Unsupervised learning of image manifolds by semidefinite programming, International Journal of Computer Vision 70 (1) (2006) 77–90.
[7] Y.W. Teh, S. Roweis, Automatic alignment of local representations, in: Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, 2003, pp. 841–848.
[8] X. He, D. Cai, S. Yan, H. Zhang, Neighborhood preserving embedding, in: IEEE International Conference on Computer Vision, Vancouver, Canada, 2005, pp. 1208–1213.
[9] X. He, P. Niyogi, Locality preserving projections, in: Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, 2004, pp. 37–44.
[10] H.J. Yin, W.L. Huang, Adaptive nonlinear manifolds and their applications to pattern recognition, Information Sciences 180 (2010) 2649–2662.
[11] H.J. Yin, On multidimensional scaling and the embedding of self-organising maps, Neural Networks 21 (2008) 160–169.
[12] S. Lafon, A.B. Lee, Diffusion maps and coarse-graining: a unified framework for dimensionality reduction, graph partitioning, and data set parameterization, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (9) (2006) 1393–1403.
[13] J.A. Lee, A. Lendasse, M. Verleysen, Nonlinear projection with curvilinear distances: isomap versus curvilinear distance analysis, Neurocomputing 57 (2004) 49–76.
[14] L. Chen, Local multidimensional scaling for nonlinear dimension reduction, graph layout and proximity analysis, Ph.D. Thesis, University of Pennsylvania, 2006.
[15] J. Lee, M. Verleysen, Nonlinear Dimensionality Reduction, Springer, Berlin, 2007.
[16] J. Venna, S. Kaski, Local multidimensional scaling, Neural Networks 19 (6–7) (2006) 889–899.
[17] M. Bernstein, V. de Silva, J.C. Langford, J.B. Tenenbaum, Graph approximations to geodesics on embedded manifolds, Technical Report, Department of Psychology, Stanford University, 2000.

[18] M. Balasubramanian, E.L. Schwartz, J.B. Tenenbaum, V. de Silva, J.C. Langford, The Isomap algorithm and topological stability, Science 295 (5552) (2002) 7.
[19] L.V.D. Maaten, E. Postma, H.V.D. Herik, Dimensionality reduction: a comparative review, Tilburg University, Technical Report, TiCC-TR 2009-005, 2009.
[20] O. Kouropteva, O. Okun, M. Pietikäinen, Incremental locally linear embedding, Pattern Recognition 38 (2005) 1764–1767.
[21] J. Lee, M. Verleysen, Quality assessment of dimensionality reduction: rank-based criteria, Neurocomputing 72 (7–9) (2009) 1431–1443.
[22] L. Chen, A. Buja, Local multidimensional scaling for nonlinear dimension reduction, graph drawing, and proximity analysis, Journal of the American Statistical Association 104 (485) (2009) 209–219.
[23] S. Kaski, J. Nikkilä, M. Oja, J. Venna, P. Törönen, E. Castrén, Trustworthiness and metrics in visualizing similarity of gene expression, BMC Bioinformatics 4 (48) (2003) 1–13.
[24] J. Lee, M. Verleysen, Rank-based quality assessment of nonlinear dimensionality reduction, in: Proceedings of 16th European Symposium on Artificial Neural Networks, 2008, pp. 49–54.
[25] D.E. Knuth, Sorting and Searching, The Art of Computer Programming, vol. 3, Addison-Wesley, Reading, MA, 1973.
[26] C. Spearman, The proof and measurement of association between two things, American Journal of Psychology 15 (1904) 72–101.
[27] T. Lin, H. Zha, Riemannian manifold learning, IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (5) (2008) 796–809.
[28] J. Wang, Z. Zhang, H. Zha, Adaptive manifold learning, in: Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, 2005, pp. 1473–1480.
[29] I.T. Jolliff, Principal Component Analysis, Springer-Verlag, New York, 1989.
[30] T. Cox, M. Cox, Multidimensional Scaling, Chapman & Hall, London, 1994.
[31] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley, New York, 2001.

Deyu Meng received the B.Sc., M.Sc., and Ph.D. degrees in 2001, 2004, and 2008, respectively, from Xi'an Jiaotong University, Xi'an, China. He is currently a lecturer with the Institute for Information and System Sciences, Faculty of Science, Xi'an Jiaotong University. His current research interests include principal component analysis, nonlinear dimensionality reduction, feature extraction and selection, compressed sensing, and sparse machine learning methods.

Yee Leung received the B.Sc. degree in geography from The Chinese University of Hong Kong in 1972, the M.Sc. and Ph.D. degrees in geography and the MS degree in engineering from The University of Colorado, in 1974, 1977, and 1977, respectively. He is currently a professor of geography of the Department of Geography and Resource Management, The Chinese University of Hong Kong. His current research interests include specialization cover the development and application of intelligent spatial decision support systems, spatial optimization, fuzzy sets and logic, neural networks, and evolutionary computation.

Zongben Xu received the M.Sc. degree in mathematics and the Ph.D. degree in applied mathematics from Xi'an Jiaotong University, Xi'an, China, in 1981 and 1987, respectively. In 1988, he was a postdoctoral researcher with the department of mathematics, The University of Strathclyde, Glasgow, U.K. He was a research fellow with the information engineering department from February 1992 to March 1994, the center for environmental studies from April 1995 to August 1995, and the mechanical engineering and automation department from September 1996 to October 1996, The Chinese University of Hong Kong, Shatin, Hong Kong. From January 1995 to April 1995, he was a research fellow with the department of computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong. He is currently a professor with the Information and System Sciences, Faculty of Science, Xi'an Jiaotong University. His current research interests include machine learning, neural networks, evolutionary computation, and multiple-objective decision-making theory.