Identification of Spinal Deformity Classification With Total Curvature Analysis and Artificial Neural Network

Hong Lin

***Abstract*—In this paper, a multilayer feed-forward, back-propagation (MLFF/BP) artificial neural network (ANN) was implemented to identify the classification patterns of the scoliosis spinal deformity. At the first step, the simplified 3-D spine model was constructed based on the coronal and sagittal X-ray images. The features of the central axis curve of the spinal deformity patterns in 3-D space were extracted by the total curvature analysis. The discrete form of the total curvature, including the curvature and the torsion of the central axis of the simplified 3-D spine model was derived from the difference quotients. The total curvature values of 17 vertebrae from the first thoracic to the fifth lumbar spine formed a Euclidean space of 17 dimensions. The King classification model was tested on this MLFF/BP ANN identification system. The 17 total curvature values were presented to the input layer of MLFF/BP ANN. In the output layer there were five neurons representing five King classification types. A total of 37 spinal deformity patterns from scoliosis patients were selected. These 37 patterns were divided into two groups. The training group had 25 patterns and testing group had 12 patterns. The 25-pattern training group was further divided into five subsets. Based on the definition of King classification system, each subset contained all five King types. The network training was conducted on these five subsets by the hold-out method, one of cross-validation variants, and the early stop method. In each one of the five cross-validation sessions, four subsets were alternatively used for estimation learning and one subset left was used for validation learning. Final network testing was conducted with remaining 12 patterns in testing group after the MLFF/BP ANN was trained by all five subsets in training group. The performance of the neural network was evaluated by comparing between two network topologies, one with one hidden layer and another with two hidden layers. The results are shown in three tables. The first table shows network errors in estimation learning and the second table shows identification rates in validation learning. The network errors and identification rates in the last round of network training and testing are shown in the third table. Each table has a comparison for both one hidden layer and two hidden layer networks.**

***Index Terms*—Artificial neural network (ANN), difference quotients, scoliosis, space curve, spinal deformity classification, torsion, total curvature analysis.**

I. Introduction

Identification of scoliosis deformity classification plays an important role for the preoperative surgical planning. By identifying the classification patterns of the patient's spinal deformity, the segmental instrumentation can be applied on the suggested levels of the scoliosis spine for the deformity corrections. There are currently two widely recognized classification methods, King classification and Lenke classification. King classification has five patterns and the Lenke classification, in theory, has 42 potential patterns in three categories [1], [2]. Identification of the spinal deformity classification, in certain circumstances, can be ambiguous and the reliability is an issue of concern. Until recently, there are still debates focusing on which classification system is superior to the other regarding to their reliabilities among inter- and intra-observers [3]. Ogon *et al.* claimed that the Lenke classification method has overall better reliabilities [4] while Richards *et al.*, through their studies, concluded that the King classification method, in certain circumstances, has higher reliability than that of Lenke [5]. The traditional way of identifying the spinal deformity classification is made by measuring the angles of scoliosis curves (Cobb or Ferguson angle) directly on the X-ray films [6]. In recent years, the computer programs have been developed to obtain the angles of scoliosis curves or to extract the features of spinal deformity from digital X-ray images so as to identify the deformity classification automatically [7]–[9]. Stokes has been working on the rule-based computer assisted King classification system until recently [7], [8]. Their system starts with on screen digitizing the four corners of each individual vertebra from coronal radiograph images. Upon the Cobb angles of each deformed curve is obtained based on these digitized points, the rule-based algorithm is used to identify the King deformity classification types. In 2004, the author, Lin, reported that they developed a computer assisted scoliosis deformity classification system to identify the Lenke spinal deformity classification [9]. This automated classification system takes two orthogonal radiograph images from coronal and sagittal planes first. Then 3-D Bezier curves are fitted onto the central axis of scoliosis spine in both planes and a simplified 3-D spine model is constructed for each scoliosis patient. By applying the differential geometry algorithm onto the central axis of the 3-D spine model, neutral, and apex of the spinal curves can be located. Upon determining the magnitude of each curve in three spinal regions such as proximal thoracic, main thoracic and thoracolumbar both in coronal and sagittal planes, Lenke classification types can be identified by a different rule-based algorithm as well. In 2006, Duong *et al.* described their scoliosis classification system by applying the fuzzy clustering algorithm [10]. In order to obtain the 3-D features of spinal deformity, this system transforms the complex 3-D spinal structure into a simplified representation, i.e., a 3-D curve passing through the center of each vertebra to represent the deformed spine. By adapting an unsupervised learning algorithm, this system applies the fuzzy K-mean clustering algorithm to regroup those space curves based on the feature similarities. In other words, this fuzzy clustering system is creating a brand new spinal deformity classification system which is necessary for the clinicians to get used to in the clinical applications and environments. By taking a different approach in this paper, Lin aimed at identifying the recognized existing classification systems automatically by adapting the traditional multilayer feed-forward, back-propagation (MLFF/BP) artificial neural network (ANN), a supervised learning algorithm. Similar to the previously mentioned two-class identification systems, this MLFF/BP ANN identification system is also based on the 3-D space curve from simplified spine model. Upon the construction of the simplified spine model, the features of the space curve, i.e., the total curvature of the central axis of 3-D spine were extracted. The total curvature, which is the sum of the square of curvature and torsion, represents the property of the space curve, in this case, the property of the central axis of 3-D spine [11], [12]. The curvature measures how sharply a curve is turning and the torsion measures its twist in 3-D space. Following the second step, the feature properties of total curvature extracted were presented onto the input layer of the MLFF/BP ANN. Instead of analyzing the torsion properties of the parametric space curve for the 3-D spine model, as Poncet *et al.* described in their literature [13], the discrete form of total curvature was introduced in this study. The total curvature values of 17 vertebrae from the first thoracic to the fifth lumbar spine obtained from the simplified 3-D spine model formed an input vector of 17 dimensions in the Euclidean space. In the network training stage, these input vectors were specified with known King classification types. After the training was fine tuned, i.e., the error value and the number of epochs reached a satisfied level, the weights were saved.

Manuscript received May 2, 2006; revised December 26, 2006.

The author is with the Seay Center for Musculoskeletal Research, Texas Scottish Rite Hospital for Children, 2222 Welborn Street, Dallas, TX 75252 USA (e-mail: honglin999@ ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2007.894831

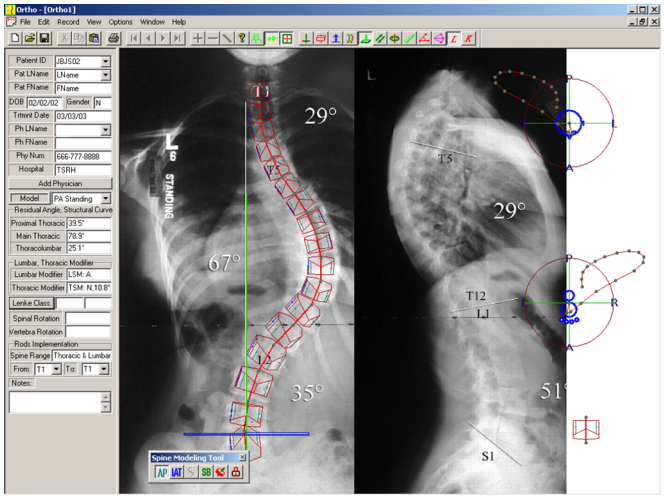

Fig. 1. Simplified 3-D spine model constructed based on coronal and sagittal spinal images.

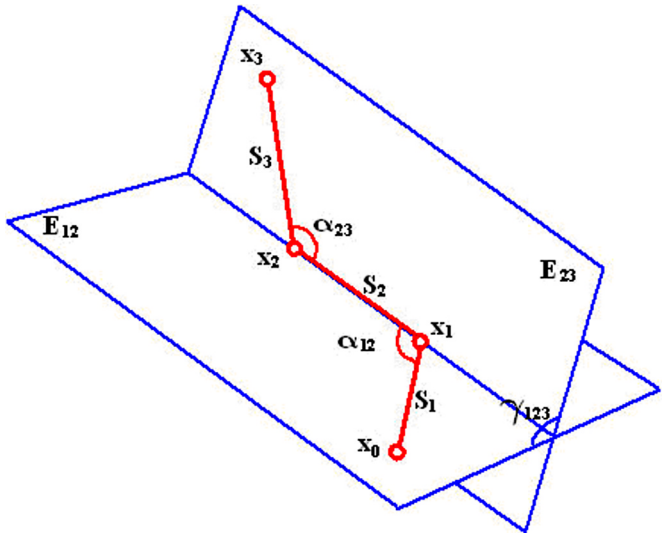

Fig. 2. Geometrical representation of total curvature.

With the saved weights, the protocol input vectors, or, in this case, the total curvature of spinal deformity patterns to be identified were fed into the input layer and this trained MLFF/BP ANN was able to identify the classification of unspecified patterns of the scoliosis spine models.

II. Methodologies

A. Simplified 3-D Spine Model

Two orthogonal X-ray images taken from coronal and sagittal planes were acquired. Presumably there was no movement for the scoliosis patient at the time when the X-rays were taking. These two images were then retrieved and were displayed on the computer screen by the modeling program [14]. Three 3-D Bezier curves with total of 17 uniformed segments and 18 nodes were superimposed interactively over the X-ray images. Each node on the curve were fitted onto the longitudinal central axis of spine and inter-vertebral disk found in both coronal and sagittal plane images. These small segments on the space curve represented the longitudinal axis of each vertebra. After the 3-D Bezier curve fitting, a series of polyhedrons was proportionally implemented onto each of the small segments representing the simplified vertebra. Thus, a space curve with 17 segments centered with 17 polyhedrons formed a simplified 3-D spine model. Simultaneously, a top and a bottom axial view of this space curve were also plotted next to the simplified 3-D spine model, as illustrated in Fig. 1.

B. Total Curvature

In order to extract the feature properties of the space curve, the total curvature analysis was performed. For the parameterized space curve represented as a vector function $x(s)$ with respect to its arc length s, i.e., $x(s) = [x(s)y(s)z(s)]^{\mathrm{T}}$, the total curvature is defined as $\sqrt{\tau^2 + \kappa^2}$, where the τ is the torsion and the κ is the curvature. They are defined as

$$\text{The curvature:} \quad \kappa = |\dot{\mathbf{x}} \times \ddot{\mathbf{x}}| \tag{1}$$

$$\text{The torsion:} \quad \tau = \frac{\det(\dot{\mathbf{x}}, \ddot{\mathbf{x}},)}{\kappa^2} = \rho^2 \left| \dot{\mathbf{x}} \cdot (\ddot{\mathbf{x}} \times \dddot{\mathbf{x}}) \right|. \tag{2}$$

In this application, the 3-D curve of central axis is defined by a series of points in 3-D space. These points can form a series of connected vectors, as illustrated in Fig. 2. From (1) and (2), which are for smooth space curve, a discrete form of curvature $\mathbf{K}$, and torsion $\mathbf{T}$, can be introduced. In this such case, the derivative with respect of s in (1) and (2) can be formally replaced by difference quotients, or divided difference. For the vector $\mathbf{x}(s)_i$, $i = 1, \ldots, n$, the following equations can be derived:

$$\dot{\mathbf{x}}_1 = [\mathbf{x}_1, \mathbf{x}_0] = \frac{(\mathbf{x}_1 - \mathbf{x}_0)}{\mathbf{S}_1} = \mathbf{e}_1 \tag{3}$$

$$\begin{aligned} \ddot{\mathbf{x}}_{12} &= \mathbf{2}\,[\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_0] = \frac{\mathbf{2}\,(\dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1)}{(\mathbf{S}_1 + \mathbf{S}_2)} = \frac{\mathbf{2}(\mathbf{e}_2 - \mathbf{e}_1)}{(\mathbf{S}_1 + \mathbf{S}_2)} \\ &= \frac{\mathbf{1}}{\mathbf{S}_{12}}(\mathbf{e}_2 - \mathbf{e}_1) \end{aligned} \tag{4}$$

$$\begin{aligned} \dddot{\mathbf{x}}_{123} &= \mathbf{6}\,[\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_0] = \frac{\mathbf{3}\,(\ddot{\mathbf{x}}_{23} - \ddot{\mathbf{x}}_{12})}{(\mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3)} \\ &= \frac{\mathbf{1}}{\mathbf{S}_{123}} \left[\frac{\mathbf{1}}{\mathbf{S}_{23}}(\mathbf{e}_3 - \mathbf{e}_2) - \frac{\mathbf{1}}{\mathbf{S}_{12}}(\mathbf{e}_2 - \mathbf{e}_1) \right]. \end{aligned} \tag{5}$$

From (3)–(5), and notice $|\mathbf{e_i} \times \mathbf{e_i}| = 0$, we have discrete form of curvature $\mathbf{K}$ and torsion $\mathbf{T}$

$$\begin{aligned} \mathbf{K}_{12} &= |\dot{\mathbf{x}}_1 \times \ddot{\mathbf{x}}_{12}| = \left| \dot{\mathbf{x}}_1 \times \frac{\mathbf{2}\,(\dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1)}{\mathbf{S}_1 + \mathbf{S}_2} \right| \\ &= \frac{\mathbf{2}}{\mathbf{S}_1 + \mathbf{S}_2} |\mathbf{e}_1 \times (\mathbf{e}_2 - \mathbf{e}_1)| \\ &= \frac{\mathbf{2}}{(\mathbf{S}_1 + \mathbf{S}_2)} |\mathbf{e}_1 \times \mathbf{e}_2| = \frac{\mathbf{sin}\alpha_{12}}{\mathbf{S}_{12}} \\ \mathbf{T}_{123} &= \frac{\mathbf{det}(\dot{\mathbf{x}}_1, \ddot{\mathbf{x}}_{12}, \dddot{\mathbf{x}}_{123})}{\mathbf{K}_{12} \cdot \mathbf{K}_{23}} = \frac{|\dot{\mathbf{x}} \cdot (\ddot{\mathbf{x}} \times \dddot{\mathbf{x}})|}{\mathbf{K}_{12} \cdot \mathbf{K}_{23}} \end{aligned}$$

while

$$\begin{aligned} |\dot{\mathbf{x}} \cdot (\ddot{\mathbf{x}} \times \dddot{\mathbf{x}})| &= \frac{\mathbf{1}}{\mathbf{S}_{12} \cdot \mathbf{S}_{23} \cdot \mathbf{S}_{123}} |\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3) - \mathbf{e}_1 \cdot (\mathbf{e}_1 \times \mathbf{e}_3) \\ &\qquad - \mathbf{e}_1 \cdot (\mathbf{e}_1 \times \mathbf{e}_2)| \\ &= \frac{\mathbf{1}}{\mathbf{S}_{12} \cdot \mathbf{S}_{23} \cdot \mathbf{S}_{123}} |\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)| = \frac{\mathbf{sin}\gamma_{123}}{\mathbf{S}_{123}}. \end{aligned}$$

Here $\mathbf{S}_{12}$ is the average length of segments $\mathbf{S}_1$ and $\mathbf{S}_2$, while $\mathbf{S}_{123}$ is the average length of segments $\mathbf{S}_1$, $\mathbf{S}_2$, and $\mathbf{S}_3$.

The $\boldsymbol{\alpha}_{12}$ denotes the angle of deformity between two vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, while γ_{123} denotes the torsion angle between the two planes determined by the vectors

$$\mathbf{E}_{12} = \mathbf{e}_1 \times \mathbf{e}_2, \quad \mathbf{E}_{23} = \mathbf{e}_2 \times \mathbf{e}_3.$$

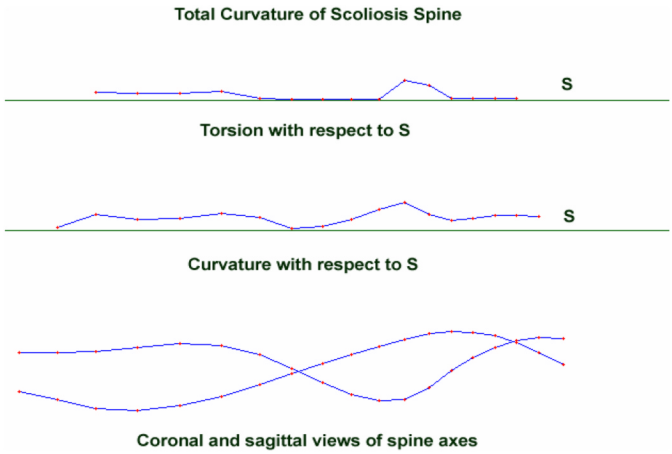

Fig. 3. Torsion and curvature with respect to S for the 3-D curve of scoliosis spine.

Fig. 3 shows both torsion and curvature values extracted from simplified spine model for King III scoliosis type.

C. Definition of King Classification

The King spinal deformity classification was defined by King [1] in 1983 for the purpose of assisting orthopedic surgeons to make decisions for the level of fusion and instrumentation on the spine of the scoliosis patient. Later, researchers have been working on a way to implement computer programs for the purpose of automating the classification process. The King classification definition can be described by the following rule-based pseudo code [7].

```
if (both thoracic and lumbar curve exist) then
   if (lumbar curve cross midline) then
      if (lumbar curve > thoracic curve) then
         Type 1 Classification.
      else
         Type 2 Classification.
      end if
   else
      Type 3 Classification.
   end if
else if (two thoracic curve and L1 tilt to 1st curve) then
   Type 5 Classification.
else if (L4 tilt to largest thoracic curve)
   Type 4 Classification.
else
   miscellaneous
end if
```

In this study, the King classification is used for the preliminary testing of the MLFF/BP ANN identification system.

D. Artificial Neural Network

A multilayer perception with MLFF/BP ANN was implemented for the purpose of automatically identifying the King classification of spinal deformities. Such network topology tends to have higher expressive power and more free parameters to fit the training data than the single layer perception which ANN does. It has an input layer, an output layer, and at least one hidden layer. In this study, the input layer had 17 neurons accepting the total curvature values of vertebrae from the first thoracic to fifth lumbar spine. There were five neurons in the output layer. These five neurons in output layer were representing the five King classification types. The hidden layers also had 17 neurons. They increased complexity of the network and were used to calculate and modify the weights. Because the miscellaneous type, as indicated in the previously shown pseudo code, rarely occurs, and those five King classification types have covered most of the spinal deformity patterns, the number of neurons in output layer had been chosen to be five instead of six so as to simplify the network structure. In a real situation, the miscellaneous type would be identified into one of the five King classification types, as closely as possible by its feature characters. The MLFF/BP ANN is called supervised network because for each training dataset, its input pattern's classification needs to be specified for the training. When it is activated, the network is going through a feed-forward, back-propagation algorithm. This algorithm has two passes for each epoch. The forward pass propagates the input pattern data from the input layer all the way to the output layer via hidden layers. The weight matrix links the two adjacent layers. When data are passing through from layer to layer, they are altered by the weight matrix. Upon the data reach the output layer the error is calculated (the differences between the actual output and the expected output). In the backward pass, the output error is used to update weights to the output layer and then the signal propagates backwards while the synaptic weights on the hidden layers are being adjusted according to the error-correcting rule. Then the next epoch is started until the output error reaches the preset value. The following pseudo-code describes the feed-forward, back-propagation algorithm [15].

```
Initialize network (inputs, weights, error)
While (error > Critical Error)
{ propagate signals forward layer by layer
     through weight matrices;
  calculate differences between expected and
     actual values in output layer;
  propagate signals backward by computing errors
     of each hidden layers from next layer;
  update all weights by using the error-correcting
     rule;
}
```

The previous process is called "generalization." The network is said to "generalize" well when the input and output mapping computed by the network is correct for testing data which are never used for creating and training the network [16]. It will produce a correct input-output mapping even when the input is slightly different from the patterns used for the network training. However, the neural network also tends to memorize the feature that is presented in the training patterns but not true of the underlying function that is to be modeled. Such phenomenon is referred to as "over-fitting" or "over-training." When this happened, the neural network loses the ability to identify the patterns which is supposed to be. This can be solved by the "cross-validation" and "early stop" methods. With these methods, the training process is being controlled either by setting the error value first, then to run the network until this error value is reached, or by giving the number of epochs before it starts running in order to investigate the error level when it stops [17]. In determining the number of hidden layers, the performance comparison was conducted between one and two hidden layer networks. As Heykin mentioned in his book [16], "A single hidden layer is sufficient for a multilayer perceptron to compute a uniformed approximation to a given training set represented by a set of inputs and a desired output. However, that is not to say that a single hidden layer is optimum in the sense of learning time, ease of implementation and generalization."

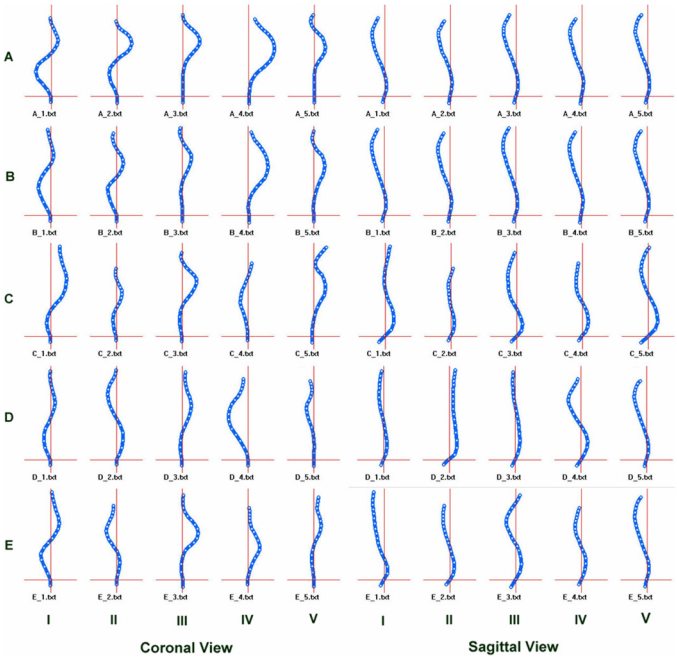

Fig. 4. Coronal view (left) and sagittal view (right) of five King classification patterns in five subsets of training group for the cross-validation method.

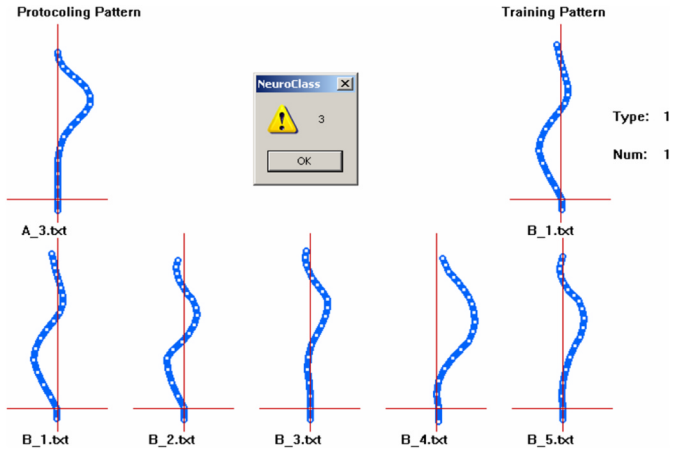

Fig. 5. Five King classification types (bottom), one of training pattern (top right), and testing pattern (top left).

E. Data Selection

A total of 37 spinal deformity patterns from scoliosis patients were selected. These 37 patterns were divided into two groups. The training group had 25 patterns and testing group had 12 patterns. The 25 patterns in training group were further divided into five subsets. For the purpose of data analysis, these five subsets were labeled A, B, C, D, and E. Based on the definition of the King classification system, each subset contained all five King types. In construction of subsets for the training group, the deformity patterns were chosen in the order of their indices and were filled up into each subset. For example, if the King type of current pattern being chosen has already existed in the subset A to be filled, this pattern would be filled into next subset B. This was continued until all five subsets were filled up with exactly five King-type patterns of scoliosis spine. As illustrated in Fig. 4, there were five rows representing five subsets in training group. Each subset in one row has five King types. The leftover patterns from training group construction were used for construction of 12 scoliosis patterns in the testing group.

F. Cross-Validation and Early Stop Methods

A MLFF/BP ANN trained with the back-propagation algorithm learns in stages, from the realization of fairly simple to more complex mapping functions as the training session progresses [16]. There are two questions arising at this point: 1) whether the MLFF/BP ANN would do better as the number of epochs for training gets larger and 2) what is the optimal number of epochs for which the network training stops? To answer these questions, it is necessary to understand that it is possible for the network to over-fit the training data if the training session is not stopped at the right point. This problem can be solved by the cross-validation and early stop methods, for which the training data are split into an estimation subset and validation subset. While training the estimation subset, the session is stopped periodically, and then the network is tested on the validation subset after each period of training in order to monitor the errors from validation subset. Because there was only a limited number of scoliosis spine model being used for this study, the hold-out method, one of cross-validation variant was implemented for network estimation and validation. The detailed implementation was that in the hold-out method, for each of the five training trials there was one subset in training group left out alternatively, leaving only four subsets in the group participated estimation learning. This left out subset was then used for validation learning. In estimation learning, the number of iterations[1] was set to be 300, 600, 900, 1200, 1500, and 1800 for the early stop strategy. Upon the estimation learning stopping at each of the previous iterations,[1] the network error was checked and recorded. Then the network was resumed with validation learning by the left out subset and number of correctly identified deformity patterns was recorded as well. These processes were conducted on two network topologies, one with one hidden layer and another with two hidden layer networks. Fig. 5 shows the user interface implementations. The bottom row illustrates five King classification patterns. The top row on the left is the unspecified pattern to be identified and the top row on the right is the training scoliosis patterns in the training group. This specified training pattern can be scrolled by replacing it with all the rest of the training patterns. After estimation and validation learning with the training group, all five subsets in the training group were participating for a last round of training. Twelve patterns in the testing group were used for testing the network. The same early stop strategy and network topologies were employed to evaluate the network's overall performances. The results are given in Section III.

III. Results

A. Simplified 3-D Spine Model and Feature Extraction

The feature extraction of Total Curvature from the central axis spine of scoliosis patients was conducted based on the simplified 3-D spine model. The modeling program provided some control points of 3-D Bezier curves for user to manipulate such curves to be fitted onto the spine center in the coronal and sagittal images. The accuracy of curve fitting was dependent on the operation of the program user. However, in normal operation, it was evaluated that in average, the error of accuracy was within 2.0 mm. This error of accuracy in modeling would have influence on the accuracy of total curvature value in feature extraction process. The simplified 3-D spinal model is illustrated in Fig. 1. The 3-D central axis representations from different scoliosis patterns are shown in Figs. 4 and 5. Based on these 3-D central axes, the features

[1]In this study, one iteration $=$ ($2\times$ number of patterns being processed) epochs. For example, in estimation learning, the number of patterns in estimation subgroup is 20. Then 100 iterations would be $100\times2\times20=4000$ epochs. However, in the final round of training, the number of patterns in the training group is 25. Then 100 iterations would be $100\times2\times25=5000$ epochs.

TABLE I
NETWORK ERROR IN ESTIMATION LEARNING

$\mathcal{E}_n$	Iterations Total	Subsets BCDE	Subsets ACDE	Subsets ABDE	Subsets ABCE	Subsets ABCD	$\overline{\mathcal{E}}_n$
1 Hidden Layer	300	0.1900	0.1000	0.1600	0.1700	0.0850	0.1410
	600	0.1100	0.0330	0.0540	0.0570	0.0260	0.0560
	900	0.0220	0.0140	0.0190	0.0200	0.0220	0.0194
	1200	0.0160	0.0100	0.0140	0.0170	0.0180	0.0150
	1500	0.0055	0.0043	0.0062	0.0096	0.0100	0.0071
	1800	0.0055	0.0037	0.0049	0.0071	0.0078	0.0058
2 Hidden Layers	300	0.2700	0.2300	0.1700	0.2000	0.1140	0.1968
	600	0.0330	0.0170	0.0250	0.2000	0.0140	0.0578
	900	0.0087	0.0089	0.0088	0.0110	0.0073	0.0089
	1200	0.0047	0.0046	0.0044	0.0054	0.0038	0.0046
	1500	0.0027	0.0026	0.0032	0.0037	0.0025	0.0029
	1800	0.0020	0.0023	0.0023	0.0031	0.0008	0.0021

$\mathcal{E}_n$ ------ Network Error from Holdout Subsets (20 patterns) in Estimation Learning.

$\overline{\mathcal{E}}_n$ ------ Average Network Error from 5 Estimation Learnings.

TABLE II
IDENTIFICATION RATE IN VALIDATION LEARNING

r_n	Iterations Total	Validate Subset A	Validate Subset B	Validate Subset C	Validate Subset D	Validate Subset E	$\overline{r}_n$
1 Hidden Layer	300	0.80	0.80	0.80	0.80	0.40	0.72
	600	0.80	1.00	0.80	0.60	0.40	0.72
	900	0.80	1.00	0.80	0.60	0.40	0.72
	1200	0.80	1.00	1.00	0.60	0.40	0.76
	1500	0.80	0.80	1.00	0.60	0.40	0.72
	1800	0.80	0.80	1.00	0.60	0.40	0.72
2 Hidden Layers	300	0.80	0.80	0.60	0.60	0.60	0.68
	600	0.80	1.00	1.00	0.60	0.60	0.80
	900	0.80	1.00	1.00	0.80	0.60	0.84
	1200	0.80	1.00	1.00	0.80	0.60	0.84
	1500	0.80	1.00	1.00	0.20	0.60	0.72
	1800	0.80	1.00	1.00	0.20	0.60	0.72

r_n ------ Identification Rate (Correctly Identified Patterns/Total Validating Patterns) in Validation Learning.

$\overline{r}_n$ ------ Average Identification Rate from 5 Validation Learnings.

of total curvature for 12 vertebrae could be extracted automatically and presented onto the input layer of MLFF/BP ANN for identification of King classification, as illustrated in Fig. 3.

B. Artificial Neural Network

Two network topologies, one with one hidden layer and another with two hidden layers were implemented for comparing the superiority between them. With cross-validation and early stop methods, the five subsets A, B, C, D, and E in training group were applied to the estimation and validation learning and the number of iterations[1] in all training trials was set to be 300, 600, 900, 1200, 1500, and 1800. The testing group of 12 patterns was used for overall network testing. In each round of training and testing, both network topologies were involved. The network error $\overline{\varepsilon}_n$ was monitored and recorded after the network estimation or training was stopped at given number of iterations.[1] After the validation or testing session the identification rate r_n based on the number of correctly identified patterns was recorded as well (instead of monitoring the network error as for estimation learning). Table I gives the network errors $\overline{\varepsilon}_n$ in estimation learning session. Five columns in Table I under Subsets BCDE through Subsets ABCD show network errors $\overline{\varepsilon}_n$ in each of five estimation learning sessions. The rows are corresponding to the number of iterations[1] for stopping the network. In each number of stopping iteration, the average network error $\overline{\varepsilon}_n$ of five different estimation sessions is calculated. Similarly, Table II shows the identification rate $\overline{r}_n$ based on the number of correctly identified patterns and their average identification rates $\overline{r}_n$ in the validation session. Following Tables I and II there are two figures plotted based on the data in these tables. The graphical representation in Fig. 6(a) indicates that the estimation learning curve of network error decreases monotonically along with an increasing number of iterations[1] in the usual manner, both in one hidden layer network and two hidden layer network. In contrast, it can be seen in Fig. 6(b) that the identification rate increases monotonically to a maximum; it then starts to decrease as the training continues. However, it is noticeable that the differences on the radius of curvatures for these two curves are distinctive. For two hidden layer network its identification rate starts from low of $r_n = 0.68$ at 300 iterations[1] and then climbs to $r_n = 0.84$ at 900-1200 iterations.[1] After reaching the maximum, it declines to $r_n = 0.72$ identification rate. For one hidden layer network, however, the curve is much flatter, ranging from low of $r_n = 0.72$ at 300 and 1800 iterations[1] to high of $r_n = 0.76$ at 1200 iterations.[1] This could raise a question on why the lower network errors in estimation learning would result in lower identification rates in validation learning beyond the 1200 iterations.[1] Experiments showed that this is due to the increasing of network errors in validation learning beyond this point [16]. This phenomenon can be explained that for the MLFF/BP ANN, we may not be able to do better by going beyond the maximum point on the validation learning session. When the network is learning beyond this point, it is essentially noise contained in the

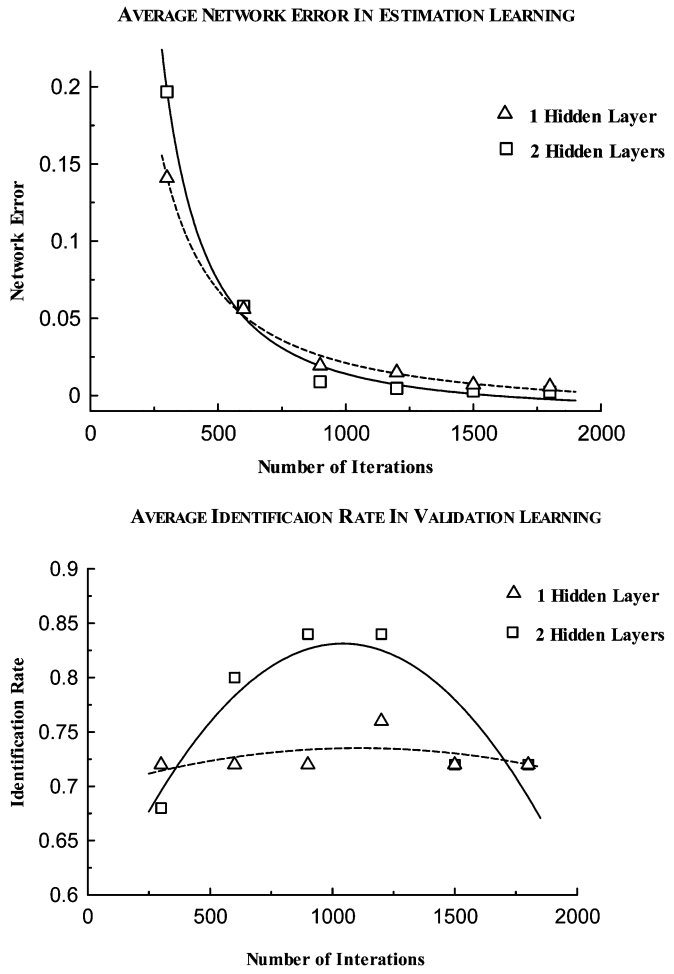

Fig. 6. (a) Average network Error in estimation learning in comparison of two network topologies for cross-validation trainings. (b) Average network identification rate in validation learning in comparison of two network topologies for cross-validation trainings.

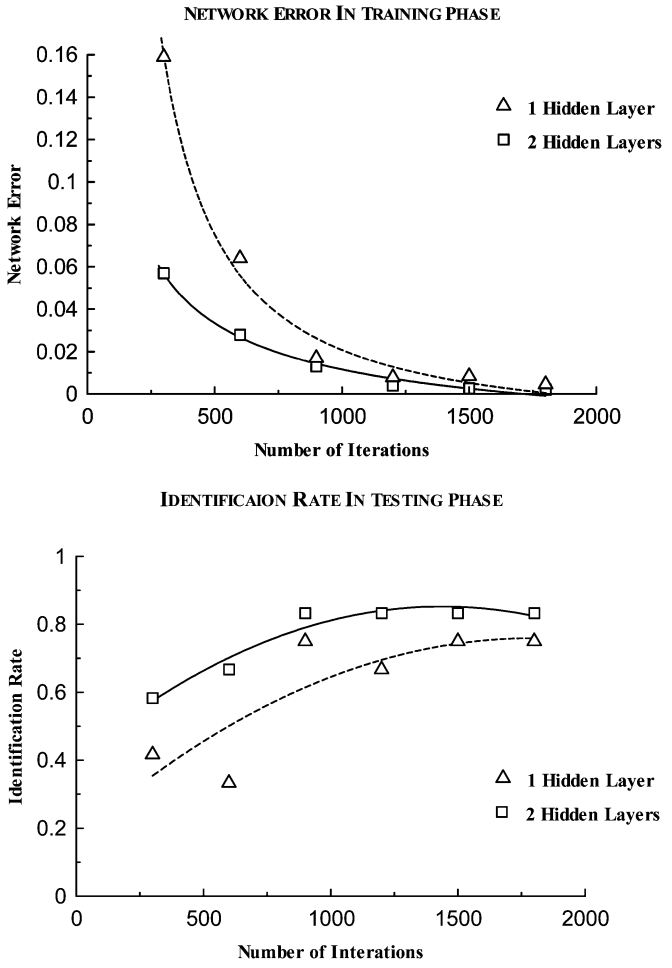

Fig. 7. (a) Network error in training session (25 patterns) in comparison of two network topologies for overall network training and testing with all 37 patterns. (b) Network identification rate in testing session (12 patterns) in comparison of two network topologies for overall network training and testing with all 37 patterns.

training data. It suggests that maximum point on the validation learning curve should be used as a sensible criterion for stopping the training session. The possible reasons for the noise existed in this training data could be explained as that there were scarcity of labeled training patterns and such extremely limited number of patterns might not be satisfactorily labeled. By comparing the results in Tables I and II as well as Fig. 6(a) and (b) between two network topologies, the conclusion may be drawn that the two hidden layer network is more sensible to the noise in the training data than one hidden layer network does. Nevertheless, the two hidden layer networks could do better if the network stopped at the right number of iterations[1] in the training session. Table III further shows the results of the last round of the training session using patterns of all five subsets in the training group, and the corresponding testing session with 12 patterns in the testing group. In Table III, the network errors and identification rates are shown for each of stopped iterations[1] from 300 to 1800 with interval of 300, both in one hidden layer and two hidden layer network. Fig. 7(a) gives the graphical representation of network errors which decreases monotonically along with an increased number of iterations[1] for both one and two hidden layer networks. This is similar to the learning curve in the estimation session. What is different is that without the cross-validation method and with all 25 patterns in the training group participated training, the identification rates increase monotonically to the maximum starting at round 1000 iterations[1] and stay as high as $r_n = 0.83$ for two hidden layer and $r_n = 0.75$ for one hidden layer networks, towards the 1800 iterations.[1] This can be seen from graphical representation in Fig. 7(b). The explanation for this could be given as that for the last round of training and testing, the number of datasets trained was 25 instead of 20, and there was no cross-validation method involved. By analyzing the data shown from Tables I–III and their plots in Figs. 6 and 7, one may reach conclusions that if the network training were stopped at appropriate number of iterations[1] so as to save the synaptic weights for identifying the King classification patterns, the two hidden layer network would do better than the one hidden layer network. This stop number should be chosen as around 1000 iterations.[1] The network would do NO better by going beyond this number.

IV. DISCUSSION

Identification of the spinal deformity classification has been an important topic in the orthopedic community. Many literatures have been published regarding their reliabilities [3]–[5]. However, in some situations, these classification methods remain ambiguous. Among the intra-observers and inter-observers, the results could be different. The purpose of this study is intended to use the ANN to reduce such ambiguity. In the results from this study, the data shown from Tables I–III do not seem to be adequate enough. This is due to the following factors: 1) the datasets used to participate study were small, only 25 patterns for

TABLE III
COMPARISON BETWEEN TWO TOPOLOGIES FOR TRAINING AND TESTING

	Iterations Total	Trained Patterns	Test Patterns	Correctly Identified	r_n	$\mathcal{E}_n$
1 Hidden Layer	300	25	12	5	0.417	0.159
	600	25	12	4	0.333	0.064
	900	25	12	9	0.750	0.017
	1200	25	12	8	0.667	0.0079
	1500	25	12	9	0.750	0.0084
	1800	25	12	9	0.750	0.0046
2 Hidden Layers	300	25	12	7	0.583	0.057
	600	25	12	8	0.667	0.028
	900	25	12	10	0.833	0.013
	1200	25	12	10	0.833	0.0040
	1500	25	12	10	0.833	0.0027
	1800	25	12	10	0.833	0.0021

$\mathcal{E}_n$ ------ Network Error with 25 Patterns in Network Training Phase.

r_n ------ Identification Rate (Correctly Identified Patterns/Total of Testing Patterns) in Network Testing Phase.

cross-validation and early stop methods of training and 12 patterns for overall network testing; 2) the classification specifications of deformity pattern were based only on the King classification literatures [1], [7] and they may be different if they were specified by different observers; 3) only eight early stop settings were performed for all trials, i.e., 300 through 1800 with an interval of 300. Nevertheless, from these tables, one can see that the two-hidden-layer network has better performance than that of one-hidden-layer network. At the same time, the results in the study confirmed the existence of over-fitting or over-training in the selected training datasets in validation learning session. With all 25 datasets trained, the two layer network could identify the testing patterns fairly well. To further improve the neural network performance, it is suggested that: 1) the number of datasets for training needs to be increased; 2) more experienced observers participate in the classification specification for the neural network training patterns; and 3) more early-stop settings need to be conducted to obtain various results from every trial in both network topologies for comparisons. However, from the author's understanding, the computerized classification systems, including this proposed neural network identification system, are not intended to replace the traditional classification method conducted by orthopedic surgeons, but for either the training purposes or assisting them for classification screening in decision making purposes.

REFERENCES

[1] H. A. King, J. H. Moe, and D. S. Bradford, "The selection of fusion levels in thoracic idiopathic scoliosis," *J. Bone Joint Surgery*, vol. 65, pp. 1302–1313, 1983.

[2] L. G. Lenke *et al.*, "Adolescent idiopathic scoliosis," *J. Bone Joint Surgery*, vol. 83, pp. 1169–1181, 2001.

[3] L. G. Lenke *et al.*, "Intraobserver and interobserver reliability of the classification of thoracic adolescent idiopathic scoliosis," *J. Bone Joint Surgery*, vol. 80a, no. 8, 1998.

[4] M. Ogon *et al.*, "Interobserver and intraobserver reliability of Lenke's new scoliosis classification system," *Spine*, vol. 27, no. 8, pp. 858–863, 2002.

[5] B. S. Richards *et al.*, "Comparison of reliability between the Lenke and King classification systems for adolescent idiopathic scoliosis using radiographs that were not pre-measured," *Spine*, vol. 28, no. 11, pp. 1148–1157, 2003.

[6] D. L. Carman, R. H. Browne, and J. G. Birch, "Measurement of scoliosis and kyphosis radiographs," *J. Bone Joint Surgery*, vol. 72, pp. 328–33, 1990.

[7] I. A. F. Stokes and D. D. Aronsson, "Identifying sources of variability in scoliosis classification using a rule-based automated algorithm," *Spine*, vol. 27, no. 24, pp. 2801–2805, 2002.

[8] I. A. F. Stokes and D. D. Aronsson, "Computer-assisted algorithms improve reliability of King classification and Cobb angle measurement of scoliosis," *Spine*, vol. 31, no. 6, pp. 665–670, 2006.

[9] H. Lin and D. J. Sucato, "Identification of Lenke spine deformity classification by simplified 3-D spine model," in *Proc. 26th Annu. Int. Conf. IEEE/EMBS*, 2004, pp. 3144–3146.

[10] L. Duong, F. Cheriet, and H. Labelle, "Three-dimensional classification of spinal deformities using fuzzy clustering," *Spine*, vol. 31, no. 8, pp. 923–930, 2006.

[11] E. Hierholzer and G. Luxmann, "Three-dimensional shape analysis of the scoliotic spine using invariant shape parameters," *J. Biomechanics*, vol. 15, no. 8, pp. 583–598, 1982.

[12] H. Lin, "Identification of spinal deformity classification with total curvature analysis and artificial neural network," in *Proc. 27th Annu. Int. Conf. IEEE/EMBS*, 2005, pp. 6168–6171.

[13] P. Poncet, J. Dansereau, and H. Labelle, "Geometric torsion in idiopathic scoliosis, three-dimensional analysis and proposal for a new classification," *Spine*, vol. 26, no. 20, pp. 2235–2243, 2001.

[14] H. Lin, "The simplified spine modeling by 3-D Bezier Curve based on the orthogonal spinal radiographic images," in *Proc. 25th Annu. Int. Conf. IEEE/EMBS*, 2003, pp. 944–946.

[15] K. Kutza, "The backpropagation network," *Neural Network at Your Fingertips*, 1996 [Online]. Available: http://www.neural-networks-at-your-fingertips.com

[16] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[17] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1986, vol. 1, pp. 316–362.

[18] I. D. Faux and M. J. Pratt, *Computation Geometry for Design and Manufacture*. Chichester, U.K.: Ellis Harwood, 1981.

[19] K. Yoshioka, "KyPlot, Version 2.0 Beta," 1997–2001.