JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 20, 1161-1182 (2004)

Modeling and Analysis of Acoustic Musical Strings Using Kelly-Lochbaum Lattice Networks

SHENG-FU LIANG AND ALVIN W. Y. SU*
Department of Electrical and Control Engineering
National Chiao Tung University
Hsinchu, 300 Taiwan
E-mail: liang@ falcon3.cn.nctu.edu.tw
**Department of Computer Science and Information Engineering*
National Cheng Kung University
Tainan, 701 Taiwan

Musical strings play important roles in the timbre of a stringed instrument. An instrument installed with different strings can produce very different sounds. To model a plucked-string instrument physically, analyzing its strings may be a good starting point. In this paper, a novel strategy for the modeling and analysis of acoustic musical strings is proposed. In order to obtain the vibration of an acoustic string, a measurement device with multiple sets of electromagnetic pickups is built. A recurrent network based on the Kelly-Lochbaum structure with multiple sets of system parameters is proposed to simulate the time-varying vibrating behavior of a target string. A fast training algorithm combining BPTT and SARPROP is also proposed to train the network automatically. The well-trained recurrent network based model can be regarded as a close simulation model of a real string, and using the model, it becomes easier to understand the characteristics of a string under different conditions when it vibrates.

Keywords: musical strings, plucked-string instruments, recurrent network, Kelly-Lochbaum structure, physical modeling

1. INTRODUCTION

A good musical instrument without high quality strings cannot produce nice sounds. Different strings can lead to large differences in timbre. The material and construction of strings are the main factors affecting how they sound. In general, good strings produce rich harmonics that usually decay slowly and, thus, are well constructed, very uniform, and close to linear. Such strings are usually used in guitars and violins. Strings used in traditional Chinese instruments are sometimes made of exotic materials, such as silk, gut, or tree fiber. They are sometimes not very uniform, and their characteristics are not linear. This also makes the sounds produced by these instruments very unique. For example, the sounds of steel strings are considered to be too hard, and the sounds of nylon strings too soft. Therefore, different constructions have been experimented with. Nylon wound steel strings and nylon wrapped silk wound steel strings are two frequently used types. They generate sounds very different from those produced by steel or nylon strings, and they

Received October 29, 2001; revised May 29, 2003; accepted September 15, 2003.
Communicated by Liang-Gee Chen.

don't sound the same as silk strings, either. In our experiment, Chin, guitar, and Pipa strings were compared. The Chin strings were nylon wrapped silk wound steel strings, the guitar strings were solid or steel wound strings, and the Pipa strings were nylon wound steel strings.

In the past few decades, numerous techniques for modeling and synthesizing strings have been developed, such as the Frequency Modulation (FM) and Wavetable techniques [1-5]. The goal of these techniques is to produce synthesized musical tones that are exactly the same as those produced by real instruments. However, they usually fall short in the following respects: sound quality, realism, complexity, flexibility, etc. Thus, physical modeling approaches have become more and more ubiquitous due to the realistic sounds they produce [6-8]. The basic idea behind the physical modeling synthesis technique is to simulate the vibration behavior of a musical instrument, for example, by modeling the reed and pipe of a wind instrument or by modeling the strings and top plate of a plucked string instrument [9-11]. Among the various physical-modeling music synthesis methods proposed so far, the Digital Waveguide Filter (DWF) based approaches are the most popular and practical ones. The general DWF for a vibrating string [7] is derived starting from the 1-D wave equation for an ideal vibrating string [12] and consists of bi-directional delay lines with loss factors within each delay line used to simulate the behavior of traveling waves propagating within a lossy vibrating string. The Kelly-Lochbaum equations [13] can be incorporated into this model to simulate the vibration-transmitting mechanism as well as the non-uniform mechanical impedance situation [14]. A model-based neural network called the Scattering Recurrent Network (SRN) was proposed to model and analyze the dynamics of various acoustic strings. A string can be approximately regarded as a one-dimensional instrument. This allows for a simpler implementation of the SRN and requires less computation to train the network. Another advantage of SRN is that the model parameters can be obtained automatically by using the measurement as the training signal and by training algorithms for the recurrent network [15]. There are also some disadvantages with of the SRN approach. First, it only works for uniform strings, such as guitar or cello strings. For Chin strings, the training seems to not be able to converge to a satisfactory solution. Second, the training takes too much time to converge. In this paper, modifications are proposed to deal with the above problems. Because it is not necessary to use a long signal to train the network, the result is a better representation of the string characteristics under certain vibration conditions obtained in a short period of time. Multiple sets of parameters can be obtained to analyze the string's characteristics under different conditions. To speed up the training and obtain better system parameters, a training algorithm that combines BPTT and SARPROP is used [15, 16]. The training data are obtained by measuring the vibration of the acoustic string excited by a single pluck. In order to obtain these measurements, we constructed a steel-string measurement system, which consists of three Dimarzio Virtual Vintage electromagnetic pickups placed in parallel and equally spaced under the string to measure the vibrations of the chosen string [17]. The measured signals from the pickups are sampled and stored in a personal computer fusing an Event Layla digital interface [18, 19]. In our experiments, the proposed recurrent network with suitable system parameters successfully simulated the vibration of a specific plucked string. If this network can be regarded as a simulation model of the target string, then it can be used to analyze the tested string because the system parameters are strongly related to the string physics. Aspects such as

the impulse response, reflection and energy absorption of strings can be easily obtained. In this study, analyses of different strings were performed using the above methodology. The differences among strings were easily revealed.

In section 2, measurements of various musical strings are presented, and their respective characteristics are discussed. In section 3, the analysis model is introduced. The relation between the proposed model and the physics of musical strings is presented. A hybrid-training algorithm and a training procedure are presented in section 4. Experiments are described in section 5. A conclusion and future works are given in section 6.

2. MEASUREMENT OF MUSICAL STRINGS

We built the string measurement device as shown in Fig. 1. A musical string is installed with its two ends fixed. The chosen target strings, Chin, guitar and Pipa strings, were all steel-based strings. Therefore, it was easier to measure their vibrations using electromagnetic pickups; hence, no attached type sensor was necessary. In this device, three Dimarzio® Virtual Vintage Solo electromagnetic pickups [17] usually found in electric guitars were placed under the string in parallel and equally spaced to measure the vibrations of the chosen steel string in different sampling positions. An electromagnetic pickup consisting of a coil with a permanent magnet was used. The plucked steel string caused magnetic flux changes for the pickups, and electrical signals were induced through the coil. The more pickup units were used in the system, the more measured data could be obtained for dynamic analysis and training of the proposed recurrent synthesis network. Furthermore, the pickup units were supplied with sliding seats so that they could be moved along the track under the string to different positions whenever required. In our experiments, the vibrations of the target strings at various sampling positions were measured synchronously. When a string was plucked, all the measurements obtained at on the pickups were sampled at a 44.1 kHz sampling rate and 16-bit quantization level, and then stored in a PC using the Event Layla digital interface [18, 19].

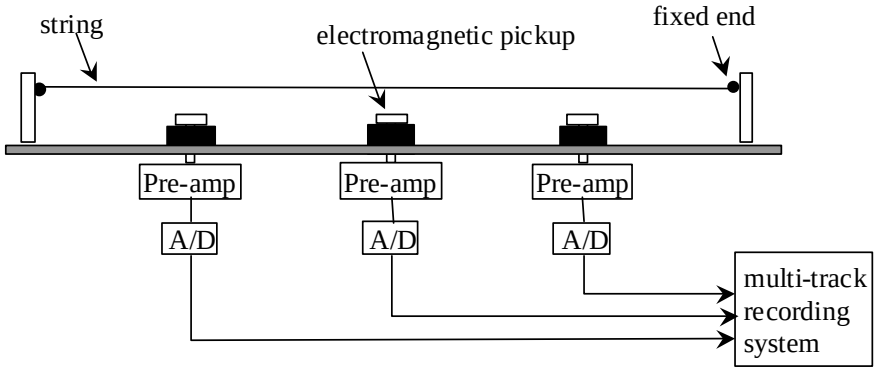
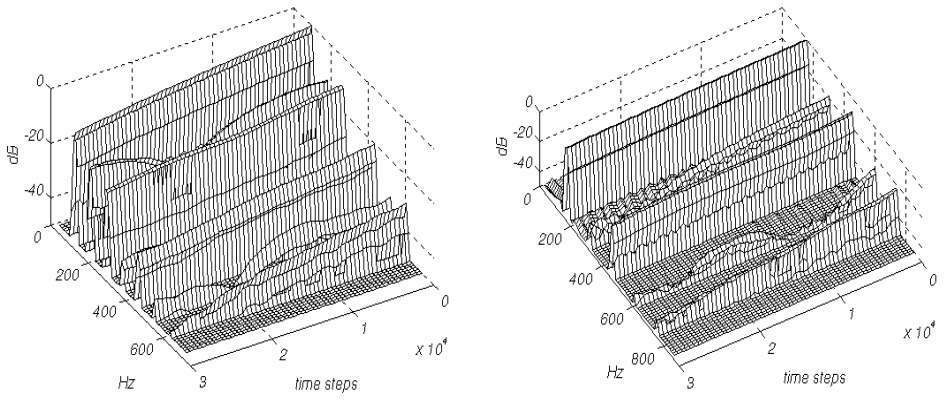


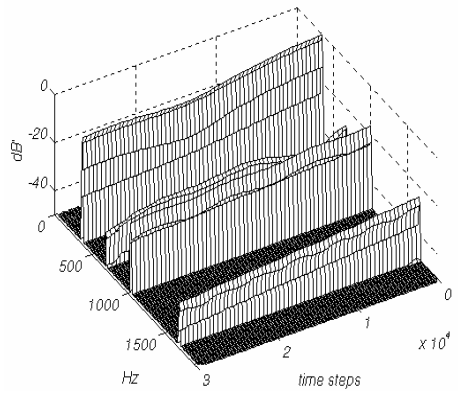
Fig. 1. The steel-string measurement set-up.

Fig. 2 shows the frequency responses of the measurements for three strings obtained through Short-Time-Fourier analyses. The harmonics and the decay pattern of the guitar string shown in Fig. 2 (c) are nice and uniform. The decay of the Pipa string is very fast as shown in Fig. 2 (b) because Pipa players usually don't want to let the instrument resonate too much due to their playing styles. The decay of the Chin string is not as fast as that of the Pipa string. Thus, the behavior of both the Pipa string and the Chin string are not as good as that of the guitar string. It was noted that it is not possible to measure a response longer than 5 seconds after plucking because the pickup was not sensitive enough and will induce more sensor noise for our analysis work.



(a)

(b)



(c)

Fig. 2. Short-time fourier transforms of three musical strings. (a) Analysis of the vibration of a Chin E2 string. (b) Analysis of the vibration of a Pipa D3 string. (c) Analysis of the vibration of a guitar E4 string.

3. SIMULATION MODEL OF MUSICAL STRINGS

In fact, an acoustic string should be considered as a 3-dimensional object, and the physical model should also be a 3-dimensional one. While it is possible to build such a model [20], too much computation may be required to obtain the model parameters. The 1-D wave equation for an ideal vibrating string was derived by Morse [12]. If an ideal string with fixed ends is plucked, it will vibrate circularly and eternally. In most situations, energy loss is caused by friction with the surrounding air, resulting in termination, and by internal friction [9, 21]. Since a real string, especially one used in a Chinese instrument, such as the Chin and Pipa, may not be uniform in its construction, scattering junctions are applied to model a non-uniform string [14]. A scattering junction deals with the situation where a traveling wave may reflect as well as pass through a position if the respective characteristic impedances from the two sides of the position are not identical [21, 22]. If there is a non-uniform junction on a string, the characteristic impedances of the two sides are Z_1 and Z_2 as shown in Fig. 3. The rightward-moving traveling wave flowing to this junction from the left-hand side and the leftward-moving traveling force wave flowing to this junction from the right-hand side are φ_r^1 and φ_l^2, respectively. Physically, the string displacement must be a continuous function of its position such that the displacement on the left-hand side of the junction is equal to the displacement on the right-hand side of this junction. In addition, the sum of the traveling waves meeting at the junction is zero [22, 23]. These two laws are identical with the situation stated by Kirchoff's node equations [24]. Let f_l^1 and f_r^2 be the traveling waves departing from the junction and flowing to the two sides, respectively. According to Kirchoff's node equations, the relation between the traveling waves can be described as follows [23]:

$$y^J = (1-\rho)\varphi_r^1 + (1+\rho)\varphi_l^2, \tag{1}$$

and

$$\begin{cases} f_l^1 = (-\rho)\varphi_r^1 + (1+\rho)\varphi_l^2 = y^J - \varphi_r^1, \\ f_r^2 = (1-\rho)\varphi_r^1 + \rho\varphi_l^2 = y^J - \varphi_l^2, \end{cases} \tag{2}$$

where ρ is the reflection coefficient defined as

$$\rho = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \tag{3}$$

and y^J represents the displacement of a junction within a plucked string. Eqs. (1) and (2), also called the *Kelly-Lochbaum* structure [23, 25], indicate that the traveling wave departing from the junction can be computed by subtracting the traveling wave belonging to the same segment flowing into the junction from the displacement magnitude of the junction. In addition, the displacement magnitude can be computed by summing the arriving traveling waves multiplied by their individual factions $(1 - \rho)$ and $(1 + \rho)$.

The *Kelly-Lochbaum Lattice structure* is regarded as a good approximation of an acoustic string. Once the computational model for a specific string with suitable

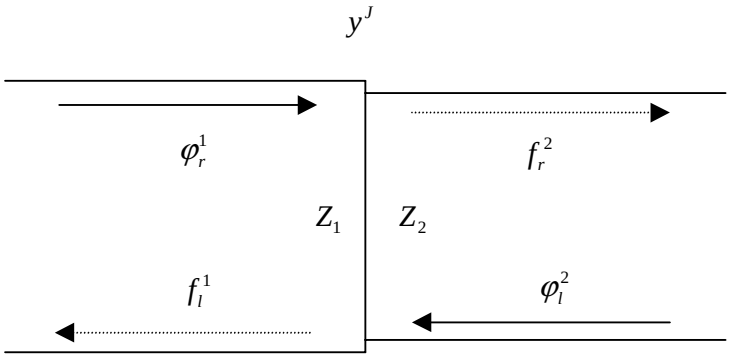


Fig. 3. A non-uniform junction, where the characteristic impedances on the two sides are not identical.

parameters is available, it is possible to accurately analyze the characteristics of the string with respect to its magnitude of vibration during vibration. To achieve the above goals, it is necessary to determine the associated parameters of the model used for a specific instrument by analyzing the recorded tones or even by modeling the instrument itself. The *Scattering Recurrent Network* (SRN) based on the *Kelly-Lochbaum Lattice structure* was proposed to model acoustic strings [23, 26]. The SRN model is shown in Fig. 4. There are two types of system parameters to be determined. Parameters of the first type, denoted by $w_{i,j}$, represent energy loss factors. Parameters of the second type, denoted by ρ_i, represent the reflection coefficient. The reflection coefficients are used to model the scattering phenomenon when a string is not uniform at a physical position. Such positions are called scattering junctions [22]. These two types of system parameters are the two main characteristics of a string. There are three types of neurons in an SRN. In Fig. 4, the big circle, labeled *y*, called the displacement neuron, represents the displacement of the simulated string at a longitudinal sampling position. The hollow circles and solid circles, labeled *f* and φ, called the departure neurons and the arrival neurons, respectively, represent the traveling wave departing from as well as flowing into a scattering junction.

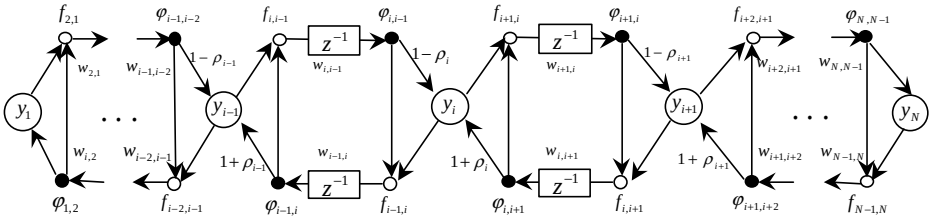


Fig. 4. The scattering recurrent network synthesis model.

There are two stages in simulating the vibration of a plucked string using the SRN model. The first stage is the *initialization* stage, and the second one is the *propagation* stage.

Initialization stage

In this stage, the initial values for displacement neurons and departure neurons are provided. Let the initial "*plucking waveform*" corresponding to displacement neuron y_i be defined as I_i. The operation of initialization stage is

for $i = 1, \ldots, N$
$y_i(0) = I_i$
$f_{i-1,i}(0) = f_{i+1,i}(0) = 0.5 \cdot y_i(0)$
end for

where N is the number of displacement neurons. Because the string is released after it is stretched and held still for a while, the initial plucking waveform is divided equally into traveling waves of the upper track and the lower track, respectively [23].

Propagation stage

After the initialization stage is finished, the computational model is ready to simulate the vibration of a plucked musical string. According to Fig. 4, the dynamics of the proposed model can be computed [25, 27]. The procedure is shown as follows:

$t = 0$
$a(\cdot)$: activation function
*net_**: net input of neuron *

Repeat
for $i = 1, 2, \ldots, N - 1$
$\varphi_{i+1,i}(t + 1) = a(net\_\varphi_{i+1,i}(t + 1)) = a(w_{i+1,i} \cdot f_{i+1,i}(t))$
$\varphi_{i,i+1}(t + 1) = a(net\_\varphi_{i,i+1}(t + 1)) = a(w_{i,i+1} \cdot f_{i,i+1}(t))$
end for

$y_1(t + 1) = y_N(t + 1) = 0$
for $i = 2, \ldots, N - 1$
$y_i(t + 1) = a(net\_y_i(t + 1)) = a((1 - \rho_i) \cdot \varphi_{i,i-1}(t + 1) + (1 + \rho_i) \cdot \varphi_{i,i+1}(t + 1))$
end for

for $i = 1, 2, \ldots, N - 1$
$f_{i+1,i}(t + 1) = a(net\_f_{i+1,i}(t + 1)) = a(y_i(t + 1) - \varphi_{i,i+1}(t + 1))$
$f_{i,i+1}(t + 1) = a(net\_f_{i,i+1}(t + 1)) = a(y_{i+1}(t + 1) - \varphi_{i+1,i}(t + 1))$
end for

$t = t + 1$
Until (synthesis stops)

It should be noted that the activation functions of neurons are all identity functions in order to match the string physics [9] and reduce the computation load. Instead of using only one set of system parameters [23], multiple sets of system parameters are used because the characteristics of a vibrating string are not fixed. This will be discussed in the

next section. The system outputs are the outputs of displacement neurons corresponding to the physical positions of pickups.

4. MULTI-STAGE TRAINING PROCEDURE

Since the string length, string tension and other factors keep changing when a string vibrates, multiple sets of system parameters have to be used. A multi-stage training procedure as shown in Fig. 5 is employed. The first set of system parameters is determined in Stage #1. The measurement data from 0 to T_1 are used as the training vector, and the resulting parameters are used to simulate the vibration in this stage. The Stage #2 training uses the measurement from T_1 to T_2 as the training vector and obtains the second set of parameters. This procedure continues until the training procedure is finished. A preset threshold is used to determine whether the difference between the simulation result and the measurement is too large and a new stage is necessary.

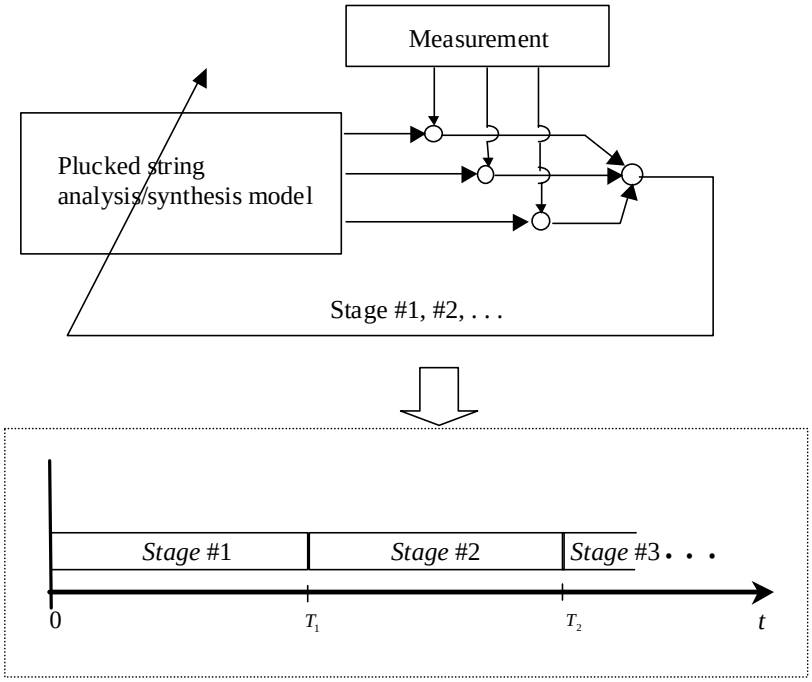


Fig. 5. Multi-stage training procedure for determining of model parameters.

For the training of recurrent networks, the backpropagation-through-time (BPTT) algorithm, which is an extension of the standard backpropagation algorithm [15, 28], is widely used. The structure of the network has to be unfolded by growing the topology layer by layer along the time steps in order to create a multi-layer feedforward network. In [23], the BPTT method was also used to train the SRN synthesis model.

In this paper, a training method that combines BPTT with a fast training algorithm, called the Simulated Annealing Resilient Back-Propagation (SARPROP) method [16], is used to train the string model, and the converging speed is much faster than that of BPTT. The SARPROP method was proposed for feedforward neural network training. This algorithm combines a quick gradient descent algorithm, called resilient backpropagation (RPROP) [29], with a simulated annealing (SA) based global search technique [30]. RPROP takes into account the sign of the gradient as seen by a particular parameter instead of the magnitude of the gradient. SA involves adding random noise to parameter updates and gradually decreasing the weights of the updates in the training process.

Though SARPROP exhibits very good performance in feedforward network training, it is not stable in recurrent network training. It has been found that the convergence of SARPROP depends very much on the initialization of training. To improve the convergence, a hybrid-training method consisting of BPTT and SARPROP is used, and the flowchart is shown as Fig. 6. In Stage #1, BPTT training is used to obtain the initial values

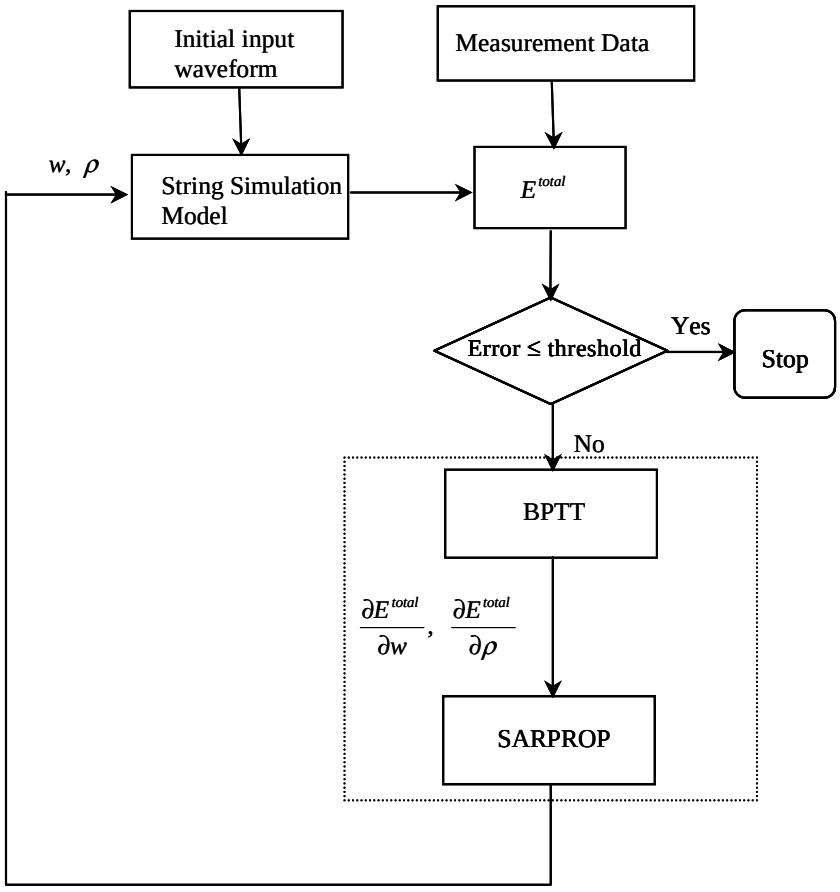


Fig. 6. The hybrid-training algorithm consisting of BPTT and SARPROP for the training procedure of the proposed model.

of the system parameters for SARPROP. Then, SARPROP is used to fine-tune these parameters to achieve stable convergence. In the other stages, BPTT is simply used to calculate the gradient values of the system parameters for each epoch. Then, they are transferred to SARPROP to determine the amount of adjustment needed for fast convergence. The device shown in Fig. 1 has three pickups. Let A be the set of displacement neurons with measured data at time t. The error signals at time t are defined as

$$e_i(t) = \begin{cases} d_i(t) - y_i(t), & i \in A, \\ 0, & \text{otherwise}, \end{cases} \tag{4}$$

where $y_i(t)$ is the output of the i-th displacement neuron at time t and $d_i(t)$ is the corresponding measurement of $y_i(t)$. The error function at time t is defined as

$$E(t) = \frac{1}{2} \sum_{i \in A(t)} e_i^2(t), \tag{5}$$

and the cost function to be minimized for one stage is

$$E^{total}(t_0, t_1) = \sum_{t=t_0+1}^{t_1} E(t), \tag{6}$$

where t_1 is the last time step and t_0 is the initial time instant in a stage. By using BPTT [18], the gradient values corresponding to the reflection coefficients can be computed as follows:

$$\frac{\partial E^{total}(t_0, t_1)}{\partial \rho_i} = \sum_{t=t_0+1}^{t_1} \delta y_i(t) \cdot (\varphi_{i,i+1}(t) - \varphi_{i,i-1}(t)), \tag{7}$$

where $\delta y_i(t)$ represents the gradient value of displacement neuron y_i at time t. The gradient values corresponding to the energy loss factors are computed as follows:

$$\begin{cases} \dfrac{\partial E^{total}(t_0, t_1)}{\partial w_{i+1,i}} = \displaystyle\sum_{t=t_0+1}^{t_1} \delta \varphi_{i+1,i}(t) \cdot f_{i+1,i}(t-1), \\ \dfrac{\partial E^{total}(t_0, t_1)}{\partial w_{i-1,i}} = \displaystyle\sum_{t=t_0+1}^{t_1} \delta \varphi_{i-1,i}(t) \cdot f_{i-1,i}(t-1), \end{cases} \tag{8}$$

where $\delta\varphi_{i+1,i}(t)$ and $\delta\varphi_{i-1,i}(t)$ represent the gradient values of arrival neurons $\varphi_{i+1,i}$ and $\varphi_{i-1,i}$ at time t, respectively. The gradient values of displacement neurons can be obtained as follows:

$$\delta y_i(t) = \frac{\partial E^{total}(t_0, t_1)}{\partial net\_y_i(t)} = (-e_i(t) + \delta f_{i-1,i}(t) + \delta f_{i+1,i}(t)) \cdot a'(net\_y_i(t)). \tag{9}$$

The gradient values of the departure neurons in the upper and lower tracks can be obtained as follows:

$$\begin{cases} \delta f_{i-1,i}(t) = \dfrac{\partial E^{total}(t_0,t_1)}{\partial net\_f_{i-1,i}(t)} = \delta\varphi_{i-1,i}(t+1)\cdot w_{i-1,i}\cdot a'(net\_f_{i-1,i}(t)), \\ \delta f_{i+1,i}(t) = \dfrac{\partial E^{total}(t_0,t_1)}{\partial net\_f_{i+1,i}(t)} = \delta\varphi_{i+1,i}(t+1)\cdot w_{i+1,i}\cdot a'(net\_f_{i+1,i}(t)). \end{cases} \tag{10}$$

Similarly, the gradient values of the arrival neurons can be obtained as follows:

$$\begin{cases} \delta\varphi_{i+1,i}(t) = \dfrac{\partial E^{total}(t_0,t_1)}{\partial net\_\varphi_{i+1,i}(t-1)} = (\delta y_{i+1}(t)\cdot(1-\rho_{i+1}) - \delta f_{i,i+1}(t))\cdot a'(net\varphi_{i+1,i}(t)), \\ \delta\varphi_{i-1,i}(t) = \dfrac{\partial E^{total}(t_0,t_1)}{\partial net\_\varphi_{i-1,i}(t-1)} = (\delta y_{i-1}(t)\cdot(1+\rho_{i-1}) - \delta f_{i,i-1}(t))\cdot a'(net\varphi_{i-1,i}(t)). \end{cases} \tag{11}$$

In Stage #1, 1000-epoch BPTT training is used to obtain the initial values of the system parameters for SARPROP. Then, 1000-epoch SARPROP (with BPTT used to calculate the gradient values of the system parameters for each epoch) is used to fine-tune these parameters to achieve stable convergence. In the other stages, BPTT is simply used to calculate the gradient values of the system parameters for each epoch. Then, they are transferred to the SARPROP to determine the amount of adjustment needed for fast convergence. The number of iterations is also 1000.

The learning parameters of SARPROP are determined empirically based on the magnitude and the size of the training signal. The initial update parameter Δ_0 is 0.0001. The temperature parameter T is 0.01. The rest of the constants are set as follows: $\eta^+ = 1.2$, $\eta^- = 0.5$, $\Delta_{max} = 0.2$, and $\Delta_{min} = 1 \times 10^{-10}$.

Instead of random values, the initial values of the synthesis parameters can be assigned as reasonable values according to the physical characteristics of the target instrument. This is different from other applications that use neural networks as their parameter-finding mechanisms. Physically, the range of the parameters corresponding to reflection coefficients should fall within [– 1, 1], and the loss factors should be very close to unity.

5. MODELING OF ACOUSTIC STRINGS

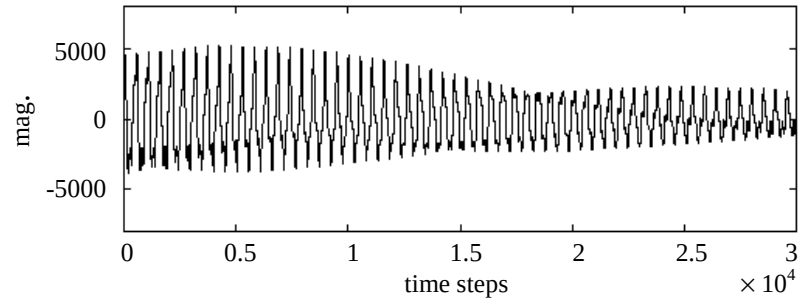
In the following experiments, we used the proposed method to model a guitar string, a Pipa string and a Chin string. The physical characteristics of the musical strings used for modeling are shown in Table 1. Three pickups were used in the device shown in Fig. 1 to measure the vibrations of the plucked strings at various positions such that the training vectors could be obtained. Pickup-1, Pickup-2 and Pickup-3 were placed in parallel and an equal distance apart under the stretched string. Pickup-1 was placed on the left side, Pickup-2 was placed in the middle and Pickup-3 was placed on the right side. In the

Table 1. Physical characteristics of the musical strings used for modeling.

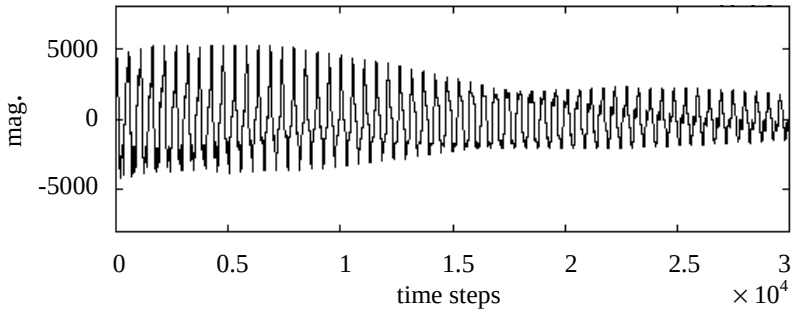
Musical strings	Chin E2 string	Pipa D3 string	Guitar E4 string
Fundamental Freq.	84 Hz	164 Hz	330 Hz
String material	Nylon-wrapped-silk-wound steel string	Nylon-wound-solid-steel string	Steel-wound-solid-steel string

training phase, the measured signals were used as the training vector. The model parameters were obtained by using the algorithm given in section 4. In the re-synthesis phase, the proposed model could simulate the vibrations of the strings. This required only a virtual pluck waveform as the excitation signal instead of a recorded wavetable of considerable size [31].

The response of the Chin E2 string measured with Pickup-3 is shown in Fig. 7 (a). The fundamental frequency was tuned to 84 Hz. The associated synthetic plucked-string tone produced by the procedure described in section 3 is shown in Fig. 7 (b). The result is very close to the original signal. The zoom-in plots of Fig. 7 are shown in Fig. 8. Figs. 8 (a) and (b) show the waveforms of the measured signal and synthetic signal of the Chin E2 string within the time steps [1, 10000], respectively. Figs. 8 (c) and (d) show the

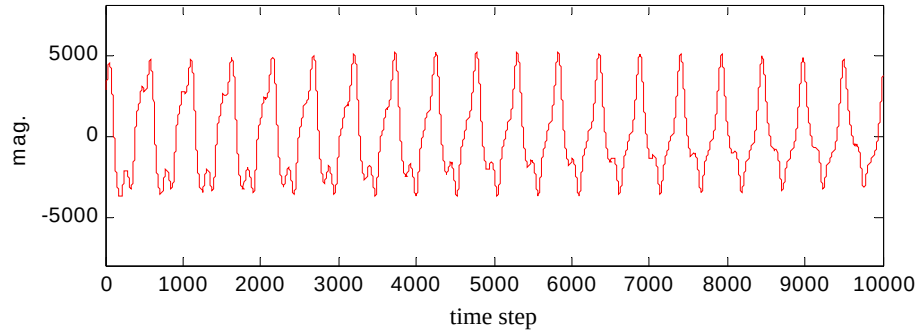


(a) Measurement.

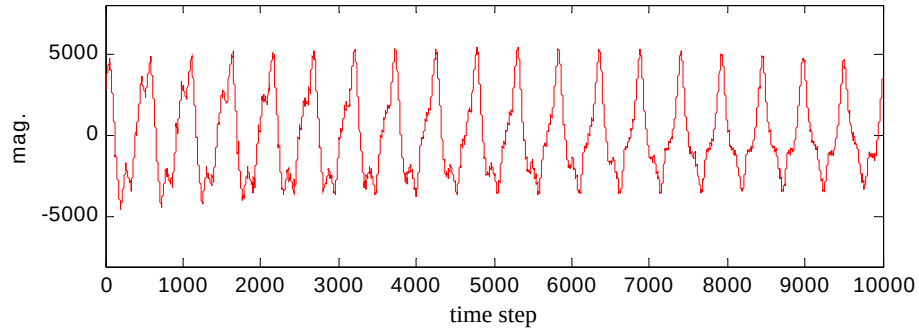


(b) Synthesis.

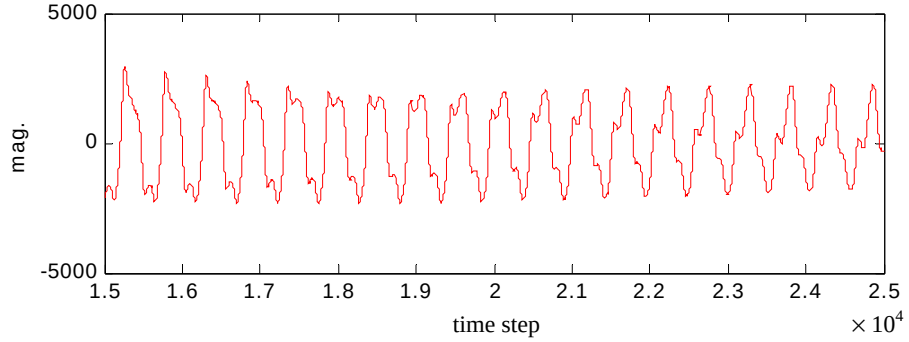
Fig. 7. Original tone and synthetic tone produced by the Chin E2 string.



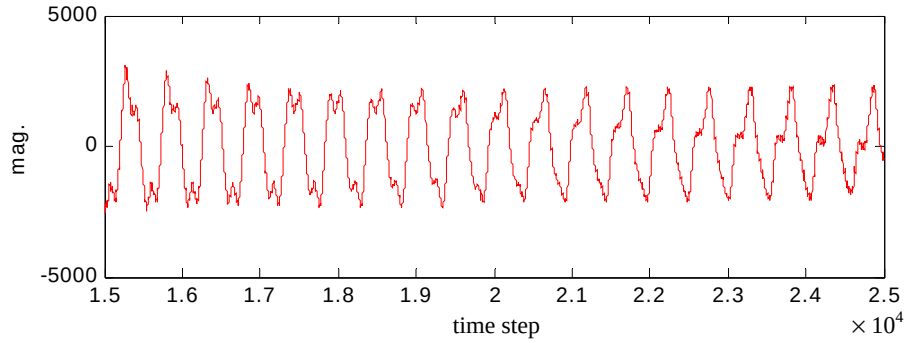
(a) Measurement.



(b) Synthesis.



(c) Measurement.



(d) Synthesis.

Fig. 8. The zoom-in plots of Fig. 7. (a) and (b) show the waveforms of the measured signal and synthetic signal of the Chin E2 string within the time steps [1, 10000], respectively. (c) and (d) show the waveforms of the measured signal and synthetic signal of the Chin E2 string within the time steps [15001, 25000], respectively.

waveforms of the measured signal and synthetic signal within the time steps [15001, 25000], respectively. We can find that the characteristics of the waveforms in these two intervals are quite different; however, the synthetic results of the proposed system agree with the measurement data pretty well. Comparing the Short-Time-Fourier-Transforms of the two signals shown in Fig. 2 (a) and Fig. 9 reveals that they are also very close to each other. In fact, if the time domain responses are similar, the frequency domain responses should also be similar since they are just different ways of representing signals. The phase response of STFT is usually not shown because human ears are insensitive to phase information when tones are nearly stable. It should be noted that phase is very important in the transient. Good time domain responses represent good magnitude responses and phase responses. Therefore, only time domain results will shown for the rest of the experiments.

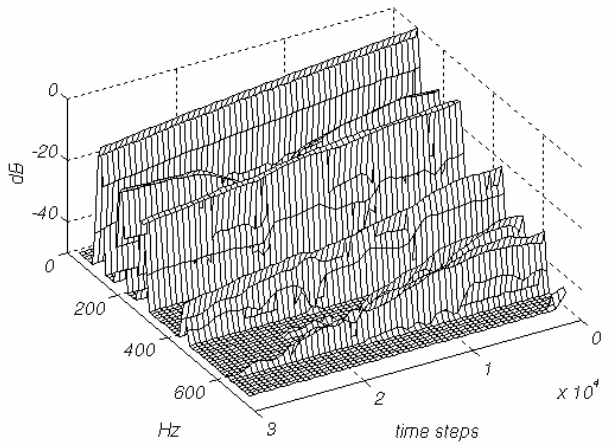
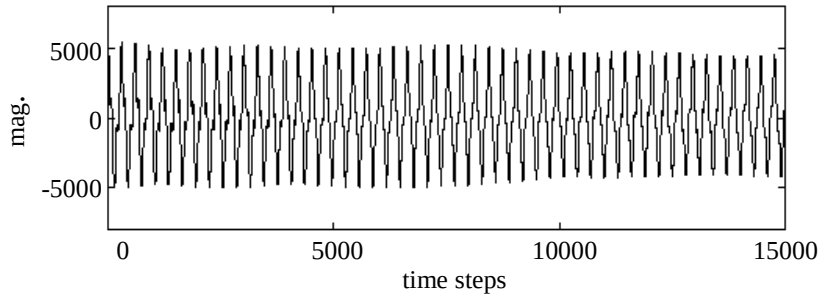
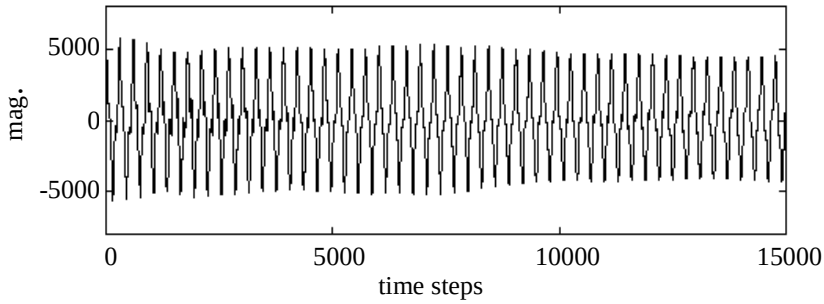


Fig. 9. STFT of the Chin E2 synthetic tone.

The Pipa D3 string was tuned to 146 Hz, and the guitar E4 string was tuned to 330 Hz. The time domain responses of the original signals and the synthetic signals for the Pipa string and the guitar string are shown in Figs. 10 and 11, respectively. The original signal and the synthetic signal are very similar in these two cases. As expected, the guitar string has the most uniform characteristics of all, and the Chin string has the most complicated response because of its unusual structure and material. The second harmonic of the Chin string decays in about the first 0.5 second after the pluck and recover afterwards. Then, the second harmonic decays with the other harmonics after they reach a peak value. This is a unique feature not found in other strings. At first, we suspected that this came from the influence of the string set-up device. After we tried all the strings (6 guitar strings, 4 Pipa strings and 7 Chin strings), we found that only the 5 Chin strings (C2, D2, E2, G2, and A2) had this kind of behavior. Nevertheless, the beating phenomenon still existed in the measured waveform [32]. This was due to the fact that the two termination ends were not actually mechanically grounded. This shouldn't be a problem once a more rigid setup device is built. The proposed methodology still works.

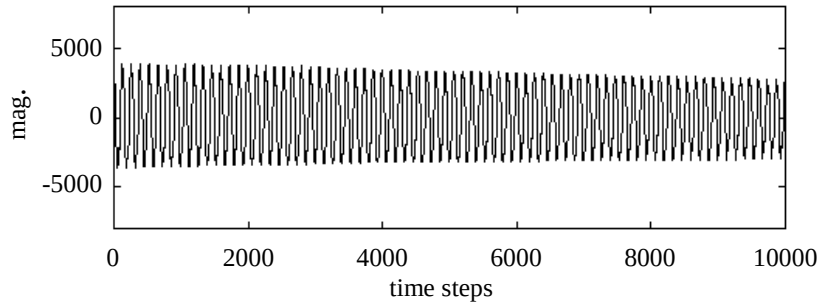


(a) Measurement.

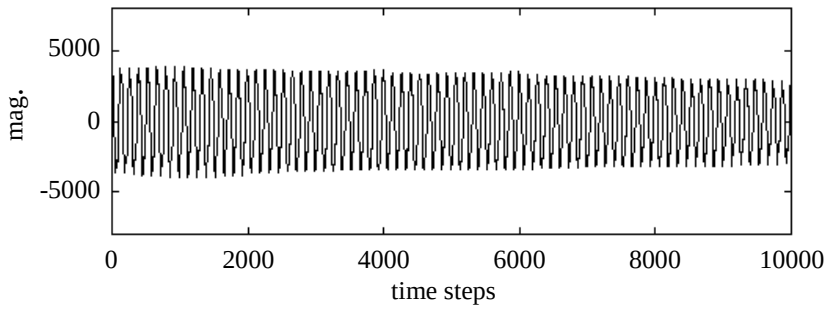


(b) Synthesis.

Fig. 10. Original tone and synthetic tone of the Pipa D3 string.



(a) Measurement.



(b) Synthesis.

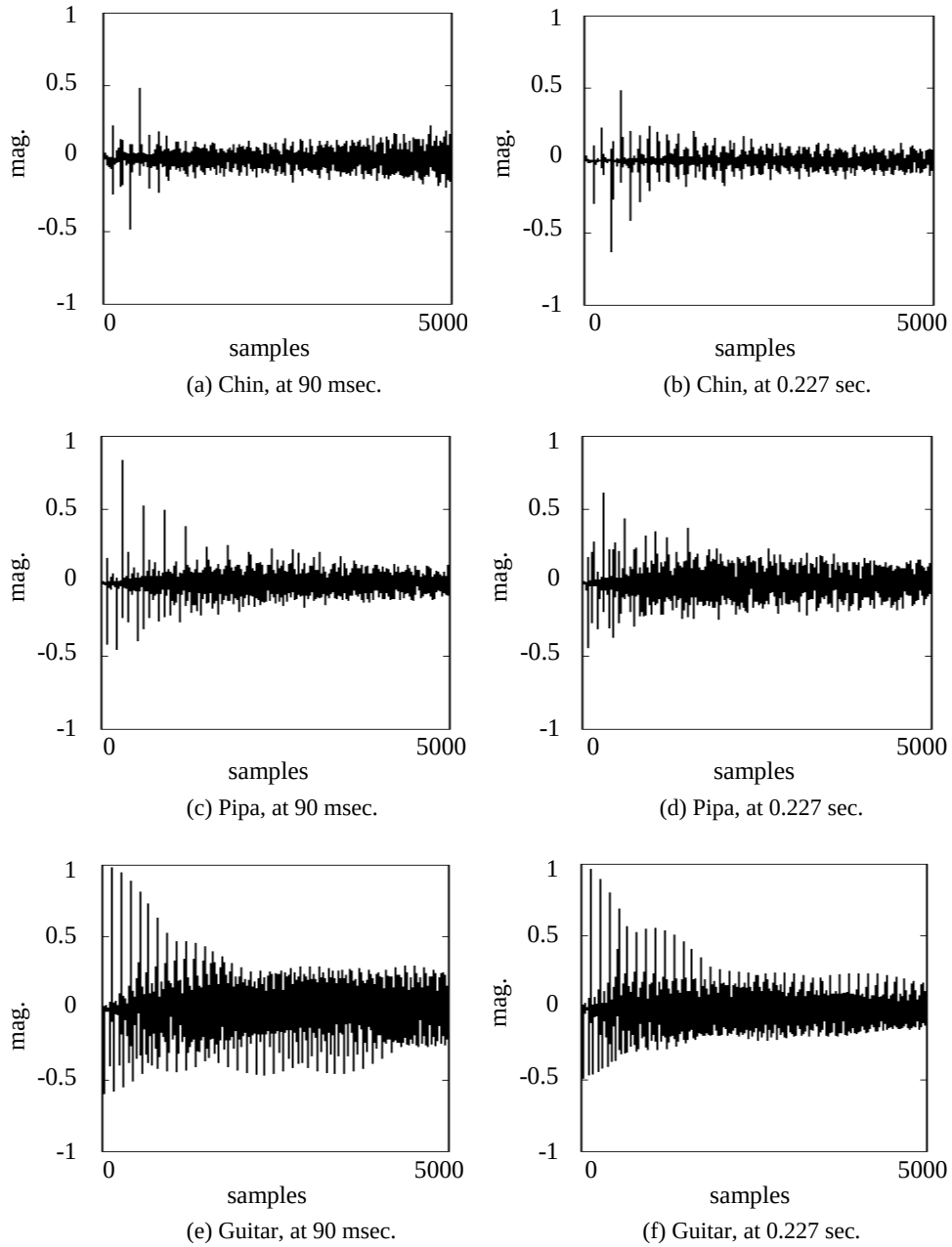
Fig. 11. Original tone and synthetic tone of the guitar E4 string.

Table 2 shows the experimental results of musical-string modeling performed using the proposed method. The Signal-To-Noise-Ratios (SNR) are 21.41 dB, 23.51 dB and 27.01 dB for the Chin, Pipa and guitar strings, respectively. Most of the errors are contributed by the high frequency harmonics because the higher harmonics have much less energy than the lower harmonics. Therefore, the training emphasizes minimization of the lower harmonics error. Once the training is trapped in a local minimum, it is very difficult to leave the minimum and search for a new direction to minimize the higher harmonics error. This is a common characteristic of most gradient type algorithms [27]. To improve the SNR, it is necessary to modify the training algorithm. Nevertheless, the proposed model is still quite accurate for lower harmonics responses that affect the timbre most.

Table 2. Experimental results of musical-string modeling conducted using the proposed method.

Musical strings	Chin E2 string	Pipa D3 string	Guitar E4 string
Numbers of iterations for training	Stage #1: BPTT: 1000 SARPROP:1000 The other stages: SARPROP:1000	Stage #1: BPTT: 1000 SARPROP:1000 The other stages: SARPROP:1000	Stage #1: BPTT: 1000 SARPROP:1000 The other stages: SARPROP:1000
Number of Stages	10	10	10
SNR	21.41 dB	23.51 dB	27.01 dB

Since the response of the simulation model is close to that of its acoustic counterpart, it is possible for us to better understand the characteristics of the target string through its associated simulation model. As stated previously, the characteristics of a string are not fixed. The simulation models of a string should also be different for different time periods. It is difficult to characterize such a changing system by using STFT analysis only because STFT analysis requires a signal segment of considerable size to maintain the frequency resolution. The proposed method can be used as a supplement since it does not require such a signal for training. Because impulse response analysis is one of the best ways to understand a system, impulse signals were applied to the simulation models corresponding to the three strings. Fig. 12 shows the impulse responses of three different simulation models at 90 msec and 0.227 sec, respectively. The frequency analyses of these three models at different time instants are shown in Fig. 13. By examining the impulse responses and frequency responses of the same string shown in Figs. 12 (a-b) and Figs. 13 (a-b), it is easier to see the difference in the string characteristics in different time periods. For example, comparing the impulse responses of the Chin string, it is found that the energy loss in the later period is larger than that in the earlier period. Another interesting phenomenon is revealed by the fact that Figs. 13 (e) and (f) are very similar to each other. That is to say, the string characteristics very little with time for the guitar string. Figs. 13 (a) and (b) are very different, which means that the characteristics change a lot. It is worth pointing out that the characteristics of the Pipa string seem to fall



(a) Chin, at 90 msec.

(b) Chin, at 0.227 sec.

(c) Pipa, at 90 msec.

(d) Pipa, at 0.227 sec.

(e) Guitar, at 90 msec.

(f) Guitar, at 0.227 sec.

Fig. 12. Impulse responses of the simulation models for the Chin E2 string, Pipa D3 string and guitar E4 string. (a), (c), and (e) were obtained by using the model parameters at 90 msec. (b), (d), and (f) were obtained by using the model parameters at 0.227 sec.

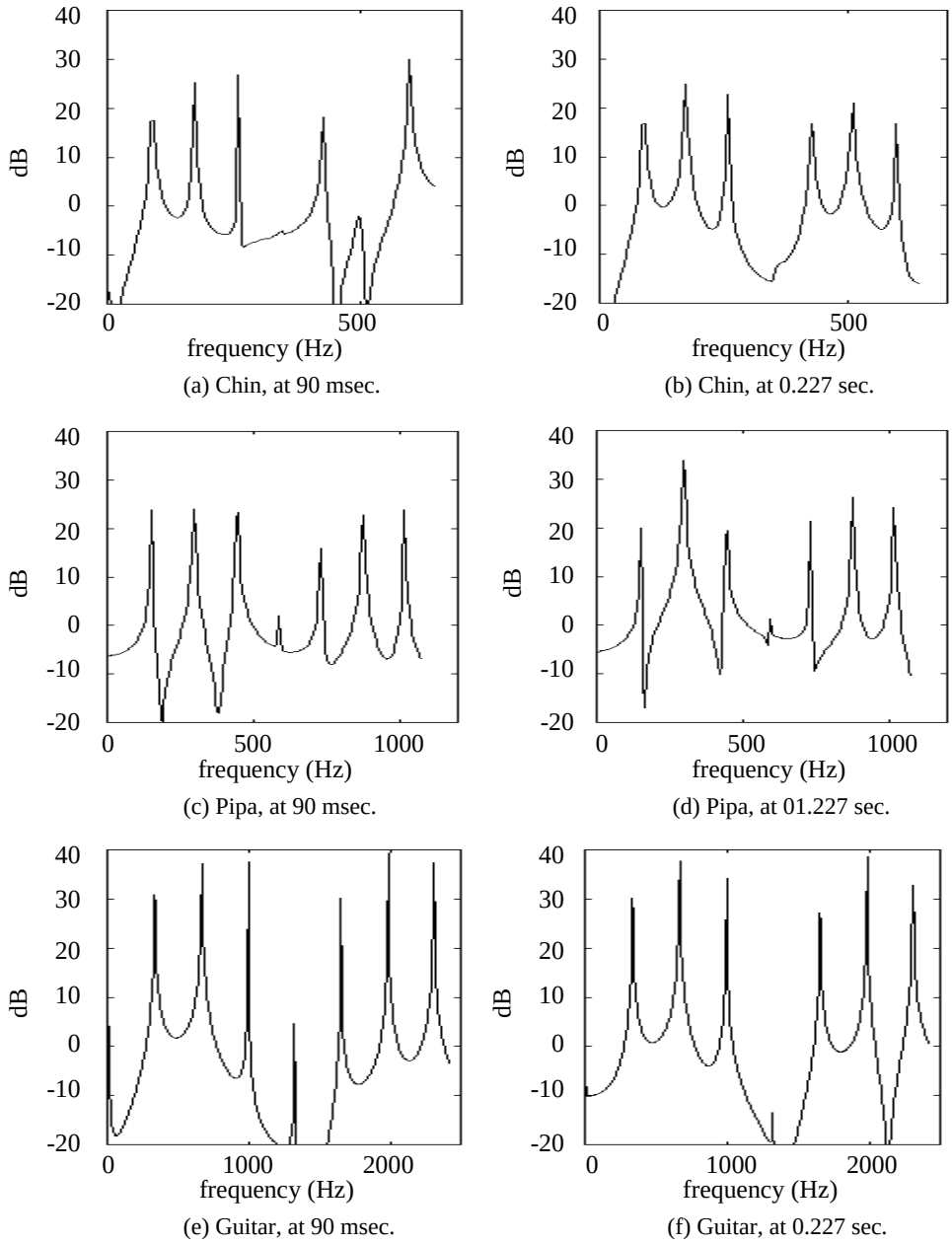


Fig. 13. The frequency analyses of the impulse responses in Fig. 12.

between those of the Chin string and the guitar string. The energy decays faster than it does for the guitar string. The Pipa string is more similar to the Chin string in this respect, based on a comparison of Figs. 12 (c-d), (a-b) and (e-f). However, the harmonics are more similar to those of the guitar string, based on a comparison of Figs. 13 (c-d), (a-b)

and (e-f). Examining Figs. 12 (a-f), it is seen that peaks appear almost periodically in the six figures. The peaks of the guitar string are very nicely distributed throughout, but those of the Chin string and the Pipa string are not as well behaved. It is not known why these phenomena happen in these strings. Perhaps string makers can enlighten us about these phenomena based on the structures and materials of strings.

To exclude the factor that the three strings were tuned to different pitches, three kinds of D3 strings were used because they are commonly used in all the instruments. The analysis results obtained at 0.3 sec are shown in Fig. 14. It is found that the total energy decay of the Pipa string is similar to that of the Chin string, but that the relationship among the lower harmonics is closer to that of the guitar string. The guitar string is a steel-wound-solid-steel string. The Chin string is a nylon-wrapped-silk-wound steel string. The Pipa string is a nylon-wound-solid-steel string without silk inside the nylon part. The Pipa string is similar to the other two strings in different respects. Whether this due to the different structures and materials of the strings or not will require further experiments and information from string makers to determine. This is certainly beyond the scope of this paper.

It should be noted that the fourth harmonics of the original signals were small because the initial pluck had less energy on this harmonic. For example, if the pluck were applied exactly in the middle position of the string, all of the even harmonics would disappear [9]. Therefore, the proposed model cannot simulate the characteristics at this harmonic. It is suggested that users pluck the string at different positions to obtain more training data.

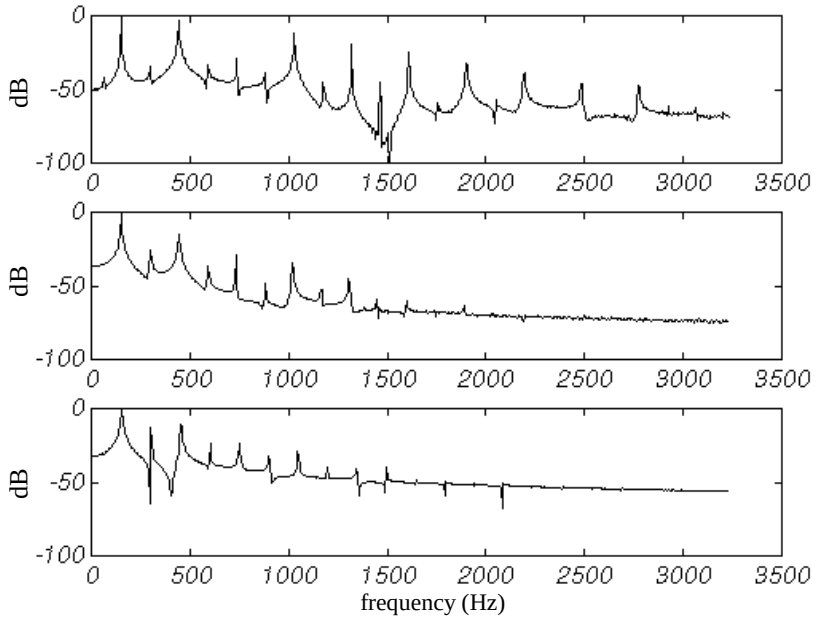


Fig. 14. The frequency analyses of the D strings of three different instruments at 0.3 sec after the pluck. The upper graph is for the guitar D string. The middle graph is for the Pipa D string. The lower graph is for the Chin D string.

6. CONCLUSIONS AND FUTURE WORKS

In this paper, a new method for analyzing acoustic strings has been presented. The structure of the simulation model has been derived and configured based on string physics. The necessary model parameters for the target strings have been obtained by using the proposed training algorithm and measurements of these plucked strings. Experiments have been performed using different kinds of strings. Combined with STFT analysis, this method provides different kinds of information about musical strings.

Another advantage of the proposed method is that it can be applied to synthesize the tones of a specific string. It is known that synthesizing guitar instruments tones may not be very difficult. However, to synthesize the tones of a specific guitar may not be easy. This research has provided a general methodology for this purpose although synthesizing a guitar tone should also involve modeling of the bridge, body, etc. if physical modeling methods are used.

The proposed method also has its limitations. For example, if the energy of some harmonics is much lower than that of others, the current training algorithm tends to neglect their contribution, and the simulation model is not able to accurately reflect the characteristics of the string at these harmonics. If this is caused by the plucking positions, plucking the string many times at different positions should solve this problem.

Our future works can be summarized as follows. A new training algorithm will be needed to improve the SNR by taking the energy of the harmonics into account. A new synthesis algorithm should also be developed not just for synthesizing a string but also for synthesizing a musical instrument.

ACKNOWLEDGMENTS

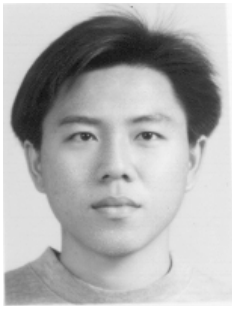
The authors are grateful for the help provided by the Pai-Ho Musical Instrument Factory, which built the string set-up device.

REFERENCES

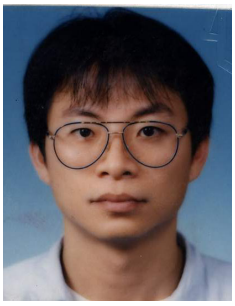
1. A. Horner, J. Beauchamp, and L. Haken, "Methods for multiple Wavetable synthesis of musical instrument tones," *Journal of the Audio Engineering Society*, Vol. 41, 1993, pp. 336-355.
2. S. D. Trautmann and N. M. Cheung, "Wavetable music synthesis for multimedia and beyond," in *Proceedings of IEEE 1st Workshop on Multimedia Signal Processing*, 1997, pp. 89-94.
3. J. M. Chowning, "The synthesis of complex audio spectra by means of frequency modulation," *Journal of the Audio Engineering Society*, Vol. 21, 1973, pp. 526-534.
4. J. M. Chowning, "Frequency modulation synthesis of the singing voices," in *Current Direction in Computer Music Research*, MIT Press, 1989,
5. B. Schottstaedt, "The simulation of natural instrument tones using frequency modulation with a complex modulating wave," *Computer Music Journal*, Vol. 1, 1977, pp. 46-50.

6. K. Karplus and A. Strong, "Digital synthesis of plucked-string and drum timbres," *Computer Music Journal*, Vol. 7, 1983, pp. 43-55.
7. J. O. Smith, "Physical modeling using digital waveguides," *Computer Music Journal*, Vol. 16, 1992, pp. 74-87.
8. J. O. Smith, "Physical modeling synthesis update," *Computer Music Journal*, Vol. 20, 1996, pp. 44-56.
9. N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, Springer-Verlag, New York, 1991.
10. G. D. Scavone, "Digital waveguide modeling of the non-linear excitation of single-reed woodwind instruments," in *Proceedings of 1995 International Computer Music Conference*, 1995, pp. 512-524.
11. S. A. V. Duyne and J. O. Smith , "Physical modeling with the 2-D digital waveguide mesh," in *Proceeding of 1993 International Computer Music Conference*, 1993, pp. 40-47.
12. P. M. Morse, *Vibration and Sound*, Published by the American Institute of Physics for the Acoustical Society of America, Woodbury, New York, 1936.
13. J. R. Deller, J. G. Proakis, and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*, Macmillan Publishing, New York, 1993.
14. S. F. Liang, "Dynamics modeling of musical string by ANN," Master thesis, Dept. of Control Engineering, National Chiao Tung University, Taiwan, 1996.
15. R. J. Williams and J. Peng, "An efficient gradient-based algorithm for on-line training of recurrent network trajectories," *Neural Computation*, Vol. 2, 1990, pp. 490-501.
16. N. K. Treadgold and T. D. Gedeon, "Simulated annealing and weight decay in adaptive learning: the SARPROP algorithm," *IEEE Transactions on Neural Networks*, Vol. 9, 1998, pp. 662-668.
17. Website of Dimarzio, http:// www.dimarzio.com/.
18. Website of Event Layla digital interface, http://www.event1.com/.
19. Website of Event Layla digital interface, http://www.soundonsound.com/sos/feb99/articles/eventlayla.694.htm/.
20. V. Duyne *et al.*, "Multidimensional digital waveguide signal synthesis system and method," U. S. Patent 5,614,686, 1997.
21. L. E. Kinsler *et al.*, *Fundamentals of Acoustics*, 3rd edition, John Wiley, 1982.
22. J. O. Smith, "Music application of digital waveguide," CCRMA Technical Report STAN-M-67, Stanford University, 1987.
23. S. F. Liang, A. W. Y. Su, and C. T. Lin, "Model-based synthesis of plucked string instruments by using a class of scattering recurrent networks," *IEEE Transactions on Neural Networks*, Vol. 11, 2000, pp.171-185.
24. L. O. Chua, C. A. Desoer, and E. S. Kuh, *Linear and Nonlinear Circuits*, McGraw-Hill, New York, 1987.
25. A. V. Oppenhim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, New Jersey, 1989.
26. W. Su and S. F. Liang, "Synthesis of plucked-string tones by physical modeling with recurrent neural networks," in *Proceedings of the IEEE 1997 Workshop on Multimedia Signal Processing*, 1997, pp. 71-76.
27. S. Haykin, *Neural Networks*, Prentice Hall, New Jersey, 1994.
28. R. J. Williams and D. Zipser, "A learning algorithm for continually running fully re-

current neural networks," *Neural Computation*, Vol. 1, 1989, pp. 270-280.
29. M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," in *Proceedings of IEEE International Conference on Neural Networks*, 1993, pp. 586-591.
30. H. Szu, "*Fast Simulated Annealing*," in J. S. Denker (ed.), *Neural Networks for Computing*, American Institute of Physics, New York, 1986, pp. 420-425.
31. V. Välimäki *et al.*, "Physical modeling of plucked-string instruments with application to real-time sound synthesis," *Journal of the Audio Engineering Society*, Vol. 44, 1996, pp. 331-353.
32. A. H. Benade, *Fundamentals of Musical Acoustics*, Dover Publication, New York, 1990.



Sheng-Fu Liang (梁勝富) was born in Tainan, Taiwan, in 1971. He received the B.S. and M.S. degrees in Control Engineering from National Chiao Tung University (NCTU), Taiwan, in 1994 and 1996, respectively. He received the Ph.D. degree in Electrical and Control Engineering from NCTU in 2000. Currently, he is a research assistant professor in Electrical and Control Engineering, NCTU. His research interests include music synthesis, neural networks, image and video processing, and brain-computer interfaces.



Alvin W. Y. Su (蘇文鈺) was born in Taiwan, 1964. He receives the B.S. degrees in Control Engineering from National Chiao Tung University in 1986. He received the M.S. and Ph.D. degrees in Electrical Engineering from Polytechnic University at Brooklyn, New York in 1990 and 1993, respectively. From 1993 to 1994, he was with Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, California. From 1994 to 1995, he was with Computer Communication Lab. of Industrial Technology Research Institute (CCL. ITRI.), Taiwan. In 1995, he joined the Department of Information Engineering and Computer Engineering at Chung Hwa University, Taiwan, where he served as an Associate Professor. In 2001, he joined the Department of Computer Science and Information Engineering, National Cheng Kung University, Taiwan. His research interests include digital audio signal processing, physical modeling of acoustic musical instruments, human computer interface design, video and color image signal processing, and VLSI signal processing. Dr. Su is a member of IEEE Computer Society and Signal Processing Society. He is also a member of Acoustical Society of America and Audio Engineering Society.