

Algorithmic approach to sonification of classical Chinese poetry

Chih-Fang Huang · Hsiang-Pin Lu · Jenny Ren

Published online: 1 September 2011
© Springer Science+Business Media, LLC 2011

Abstract The classical Chinese poetry is a remarkable form of art in traditional Chinese character. However, it is difficult for people who are unfamiliar with ancient Chinese to experience the artistic content of the poetry. In this study, a sonification scheme, *Tx2Ms (Text-to-Music)*, is proposed to extract the poetry features between lines in verses; moreover, dynamics and interval relations are modeled to map those features to the movement of multi-dimensional musical elements such as durations. This conversion is based on poetry intonation and acoustic analysis of the pronunciations of poems; and a stochastic compositional algorithm is created by applying a Markov chain. In addition, the best pentatonic mode for a specific poem is recommended according to the formants analysis. Therefore, the sonification of classical Chinese poetry not only provides a novel way for people to appreciate Chinese poetry but also enriches the state of mind and imagery in the delivery process, and the experiment results show that the proposed system is successfully accepted by most people.

Keywords Sonification · Markov chain · Classical Chinese poetry · Algorithmic composition

1 Introduction

The birth of this study is driven by the idea: “Is there any mechanism for assisting the people who are unfamiliar with ancient Chinese in experiencing the classical Chinese poetry through an alternative modality, hearing?” Fortunately, we found that both text and

C.-F. Huang (✉)
Department of Information Communication, Yuan Ze University,
135 Yuan-Tung Road, Chung-Li, Taiwan 32003, Republic of China
e-mail: jeffh@saturn.yzu.edu.tw

H.-P. Lu
Department of Mechanical Engineering, National Chiao-Tung University,
1001 University Road, Hsinchu, Taiwan 300, Republic of China

J. Ren
Institute of Music, National Chiao-Tung University,
1001 University Road, Hsinchu, Taiwan 300, Republic of China

sound can evoke emotional responses. Since each modality has its certain strengths and each combination of modalities may produce different synergistic results, sound can provide an additional and complementary perceptual channel. (Could you imagine watching a movie without any background music?) Besides, sound can be used to augment the visualization by permitting a user to visually concentrate on one field, while listening to the other. Consequently, the aim of this study is to explore the intonation and acoustic features of poems and utilize automatic generated music to strengthen the synesthesia. In other words, the features between two different kinds of medium (text and sounds) are being manipulated and mapped. Figure 1 depicts the main idea of this work, which shows the conversion from lingual text to non-speech audio, music.

The great pianist and composer in 19th century, Franz Liszt, who composed a well-know piano suite, “Petrarch Sonnet” [16], was inspired by the poetic imagery from Francesco Petraraca’s sonnet. Another great musician Gustav Mahler who composed “Das Lied von Erde” [35] in 1908, was inspired by poems in the age of Chinese Tang dynasty. This is evidence that there must be some connections between poetry and music. Although there has been research on mapping from text to sound, none is dedicated to the transformation from Chinese poetry to music. The poem-to-music conversion is a highly challenging task to any composer. It is not easy to capture the imaginary pictures of poems or to grasp the deep mind of litterateurs. This paper presents an algorithmic composition way to implement it accordingly. In this study, Chinese poems in the Five-Character-Quatrain style are used to demonstrate the proposed mapping scheme of text-to-music (Tx2Ms).

2 Related work

2.1 Sonification

The word “sonification” comprises the two Latin syllabus “sonus”, meaning sound, and the ending “fication”, forming nouns from verbs which are ending with ‘-fy’. Therefore, to “sonify” means to convey the information via sound. A Geiger detector can be seen as the very basic scientific example for sonification, which conveys (i.e., sonifies) information about the level of radiation. A clock is even more basically an example for sonification, which conveys the current time.

The word “sonification” has already been defined in a majority of researches. “*Sonification is to communicate information through nonspeech sound*” [2]; “*Sonification is the use of data to control a sound generator for the purpose of monitoring and analysis of the data*” [28]; “*Sonification is the transformation of data relations into perceived relations in an acoustic signal for the purpose of facilitating communication or*

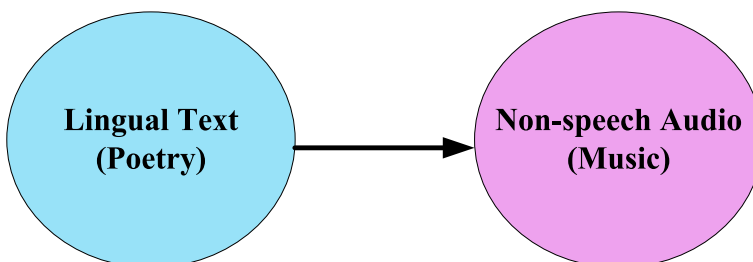


Fig. 1 Cross-disciplinary art transformation from lingual text to non-speech audio

interpretation” [30]; “Sonification is a mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purpose of interpreting, understanding, or communicating relations in the domain under study” [40].

Figure 2 illustrates the existing sonification techniques, which are already categorized into three types according to the mapping approach adopted: syntactic, semantic or lexical mapping [3, 4]. Earcons is a sound to represent an event, Auditory Icons is an auxiliary sound to complement the visual event, and Parameter Mapping serves to map the dimensional data into sound parameters. The advantage and disadvantage is also described in this figure.

Saué [39] has shown that sonification can be considered a coupling between a data set and an auditory display, and some important elements and issues should be considered in the construction of a sonification display. There are Data-centric and human-centric points of view including the functionality to be identified, tasks to be performed, and several related disciplines by Kramer [29]. It distinguishes between two major tasks in auditory displays: exploration and monitoring. In auditory display research, the possibilities of using sound in the real world are applied and extended to the human-computer interface. An auditory display supports the user in understanding any information which is represented in sound output. Brewster [5, 44] has researched on integrating non-speech audio into human-computer interface, and he showed that using musical instrument timbres (e.g. piano, trumpet) in Earcons was more effective than using simple tones (e.g. sine waves, square waves) with some applications. Goina and Polotti investigated the relationships between gesture and sound by means of an elementary gesture sonification with pitch variants [20]. Some researches use sonification techniques for mobile phone with auditory menu [48] and service notification [18]. Two practical applications are to enhance the auditory menu cues for the in-vehicle technologies [26] and to design warning signals for critical traffic situations for car drivers [14]. Signal-based sound synthesis transformed from the environment can be one of the sonification implementations [46], while MIDI-based system employs sonification to assist program understanding by rendering sonic cues [24, 47]. Although the research uses prosody specification to obtain the document-to-audio semantic

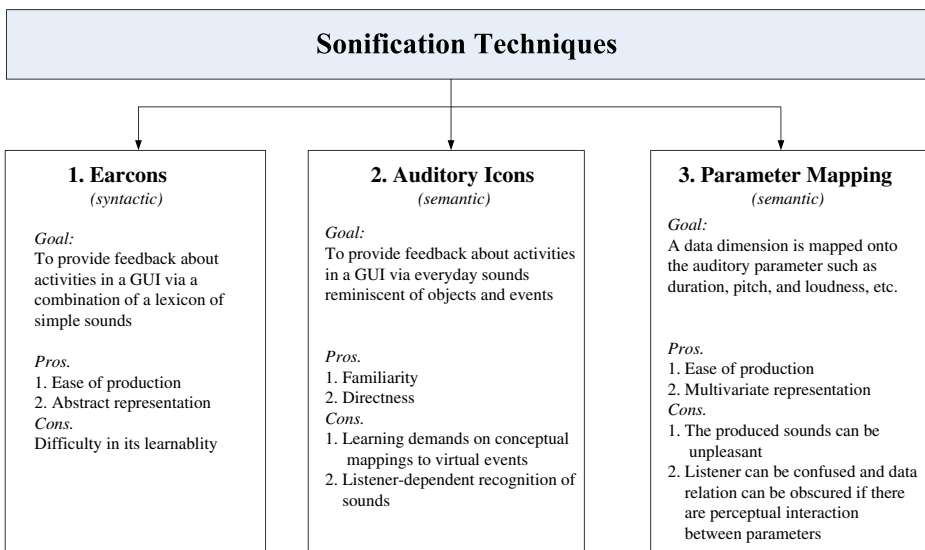


Fig. 2 Three types of existing sonification techniques. (Reorganized and Drawn based on Barrass & Kramer 1999)

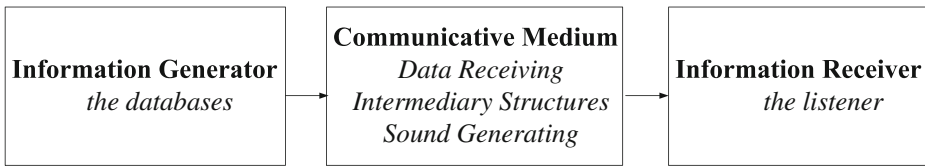


Fig. 3 Schematic of an auditory display system. *Rewritten from (Kramer, 1994a)*

correlation for sonification [42], it is mainly for the document accessibility, not for the poetry sonification, although the poetry can be analyzed and viewed as music [32].

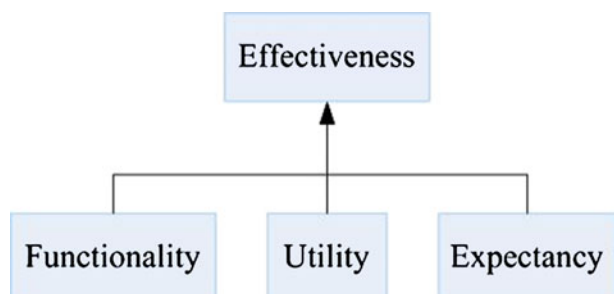
Moreover, Fig. 3 illustrates the fundamental procedure about how to design a Sonification system, where the Communicative Medium is the core of Sonification to generate sound, Information Generator is the databases to generate the top-level information, and Information Receiver is the listener to receive sound.

Although the above research shows many theories and techniques of sonification, for the poetry sonification, it needs to consider the essential goal of sonification to yield an auditory display that will be orderly and intuitively maximal in meaning (i.e., coherence) to the observer. Inevitably, the Effectiveness is what most counts in designing a sonification software or system (see Fig. 4).

- **Functionality**
 - The goal-oriented function of the system must be clearly defined.
- **Utility**
 - If the sound is ugly, people won't use it!
 - The craft of composition is important to auditory display design (i.e., a composer's skill can contribute to making auditory displays more pleasant and sonically integrated and so contribute significantly to the acceptance of such displays).
- **Expectancy**
 - Evaluation (e.g., questionnaire) is needed.

Sonification has been put into practice in a variety of areas, inclusive of medical usages, assistive technologies [5], or even data mining and information visualization. The idea of using sonification in medical usage is to use sounds to diagnose illness; the idea of carrying out sonification in assistive technologies is to make maps, diagrams and texts more accessible to the visually impaired through multimedia computer programs; the idea of applying a direct playback technique, called "Audification", in data mining and information visualization is to assist in overviewing large data sets, event recognition, signal detection,

Fig. 4 How to make an effective sonification?



model matching and education [3]. Besides, the method for rendering the complex scientific data into sounds via additive sound synthesis and further visualizing the sounds in Virtual-Reality environment has been proposed [27], which is aimed to help scientists explore and analyze huge data sets in scientific computing.

The program “Poem Generator” in Phil Winsor’s book, “*Automated Music Composition*,” has already illustrated the conversion from the constituent letters to the pitch domain by mapping their individual ASCII values onto the pitch values in MIDI. The mapping mechanism is basically derived from the idea that each character has an inherent ASCII value as its digital information in every computer. The output results convey the structure of the letters in a phrase, where rests are allocated for blank spaces and pitches are assigned for different ASCII values of the letter [51].

2.2 Classical Chinese poetry

Characters have been classified into six types by etymology: pictogram, ideogram, phonetic compounding, meaning aggregation, mixed word creation, and transliteration. Strictly speaking, the Chinese characters are logogram, primarily comprising pictograph and semasiograph, different from the phonogram, which represents phonemes (speech sounds) or combinations of phonemes. From the old days, the Chinese characters are not only a kind of symbol to record the language but also a kind of art. One of the well-known artistic creations based on the Chinese characters is classical Chinese poetry. The classical Chinese poetry uses characters for syntactic expression and for semantic narration to deliver the beauty of speech.

Jintishi, or, “modern-form poetry”, is one set of the popular poetic forms among classical Chinese poetry. In these forms, each couplet comprises a series of set tonal patterns using the four tones of the mid-ancient Chinese pronunciation. There are basically the level, rising, falling and entering tones in the classical Chinese intonation system. Furthermore, the key to the composition of Jintishi hinges on the intonation score of Ping/Ze opposition in traditional Chinese verse, where level tone belongs to Ping and the others belong to Ze. Overall, Jintishi is a specific form of classical Chinese poetry which carries consistent and well-defined rules for not only its prosody (i.e., regular meter, rhythm and intonation) but also the rhyming scheme.

Jintishi could be further categorized into three major forms based on the number of lines in each poem [22]. (All forms of Jintishi could be written in five or seven character lines.)

- Quatrain (with four lines in each poem): Some tonal patterns are followed.
- Regulated Verse (with eight lines in each poem): In addition to the tonal constraints, this form requires parallelism between the lines in the second (third and fourth lines) and third (fifth and sixth lines) couplets. The lines in these couplets have contrasting content, while the characters in each line are in the same grammatical relationship.
- Long poem in Regulated Verse (with over eight lines in each poem): This form extends the Regulated Verse to unlimited length by repeating the tonal pattern. The parallelism is required in each couplet except the first and last couplets.

However, the tonal rules received greater emphasis than parallelism. According to the tonal rules [7, 22, 53], we can infer four basic types of tonal patterns of the Five-Character Quatrain, where 1 represents Ping, and 0 represents Ze, respectively. Please refer to Table 1 for more detailed definition.

This artistic creation of Chinese characters is rich in poetic splendor, and deep in implicit imagery. The refined verses with phonemic orderliness give birth to the pleasant sounds of the recitation. It makes the poetry easy to read and to remember, which might have been lingering in the audience’s heads for days.

Table 1 Four basic types of tonal patterns of Five-Character Quatrain

Type	Name	Ping Ze structure	Ping Ze ID
I	First-line rhyming and the second syllable being Ping	PPZZP	11001
		ZZZPP	00011
		ZZPPZ	00110
		PPZZP	11001
II	First-line without rhyming and the second syllable being Ping	PPPZZ	11100
		ZZZPP	00011
		ZZPPZ	00110
		PPZZP	11001
III	First-line rhyming and the second syllable being Ze	ZZZPP	00011
		PPZZP	11001
		PPPZZ	11100
		ZZZPP	00011
IV	First-line without rhyming and the second syllable being Ze	ZZPPZ	00110
		PPZZP	11001
		PPPZZ	11100
		ZZZPP	00011

2.3 Algorithmic composition

Most of the sonification techniques use audio transform and mapping, such as the software “Coagula” [9], which uses traditional “additive synthesis” method [6] to read image data and adds up masses of sine waves, and each line in the image controls the amplitude of one oscillator at a certain pitch. The vertical position of a pixel decides the frequency, while its horizontal position corresponds to time. Another example of the image-to-sound conversion is to map the spectrogram from image to audio [25]. However this paper adopts algorithmic composition to generate music notes to integrate with a MIDI system in the music meta-level, to make the sonification more “musical”. Algorithmic composition is also named as automated composition, which uses various algorithms to compose music automatically. Computer generated algorithmic composition started in 1957, with Lejaren Hiller’s string quartet ILLIAC (Illinois Automatic Computer) Suite [21], by using a digital computer to compose music piece based on traditional rules and a Markov chain [33, 34, 45]. One of the simple approaches to performing the automatic music feature selection including pitch, chord, and others in algorithmic composition is through the mathematical function Markov chain. The input argument for this function is the current state of the musical feature, and its corresponding output, the next state, represents all the possible future musical events. Based on the transition state probability, the generative meta-level music events can be determined automatically, according to the input music data. “Sieve” algorithm [13] can be used for music generator for musical scale determination which filters the unwanted pitch classes automatically with random generator using “modulus 12”. Some other musical features (such as rhythm, dynamic, etc.) can be determined too based on the similar concept.

Several important techniques involved in the algorithmic composition such as fractal, expert system, and others have been discussed by Langston [31]. The well-known composers including Cope [10, 11] and Winsor [49, 50] introduced algorithmic composition into their pieces successfully. Cope writes some programs with algorithms not only to analyze the existing input music, but to create the new compositions to fit the original

music style. Winsor writes a practical library of compositional algorithms and related BASIC routines and C programming functions designed and tested to ease the transition into computer-generated music. Several mapping techniques for algorithmic composition, based on music between gestures or structures, have been surveyed by Doornbusch [12]. Shan and Chiu [41] discovered the rules of musical composition from given music objects to create a new music object similar to the given music objects based on the discovered composition rules.

Algorithmic Composition is also becoming one of the important sonification techniques in recent years. Some sonification methods using algorithmic composition is to sonify the scientific data, such as the research by Sturm [43] using particle physics to map the musical data to perform the algorithmic composition. The sonification technique for algorithmic composition based on image has been developed by Wu and Li [52]. Bain [1] developed the free software, Applications in Musical Sonification (AIMS), to emphasize the creative experiments using algorithmic composition to perform some mappings including historically significant compositional formalisms and mathematical models.

3 The method of automated composition based on poetry features

Throughout this section, a mechanism of mapping poetry data onto appropriate sound features is provided. The mapping problem has been regarded as the essential issue of sonification.

In Chinese poetry, we found that the poets are interested in the patterns of prosody in the composition process. In this study, the sonification mechanism of transforming classical Chinese poetry (text) into music is explored. First of all, the aesthetic features are extracted from classical Chinese poetry. Afterwards, the rules are applied to a Tx2Ms (Text-to-Music) parameter-mapping mechanism. The conversion of text-to-music mapping is based on the pronunciation properties and the syntax characteristics in classical Chinese poetry. Figure 5 illustrates a general paradigm of our sonification in this study, including data input, data analysis, data features extraction, parameter classification between music and data, algorithmic composition, and music output.

A limited number of features and corresponding sonic attributes are taken into account so as to keep the resultant sounds as simple as possible and easy to decode because the listeners always wish to hear what the data is doing. However, the sound still will be rich in itself.

3.1 Tx2Ms (text-to-music mapping of classical Chinese poetry)

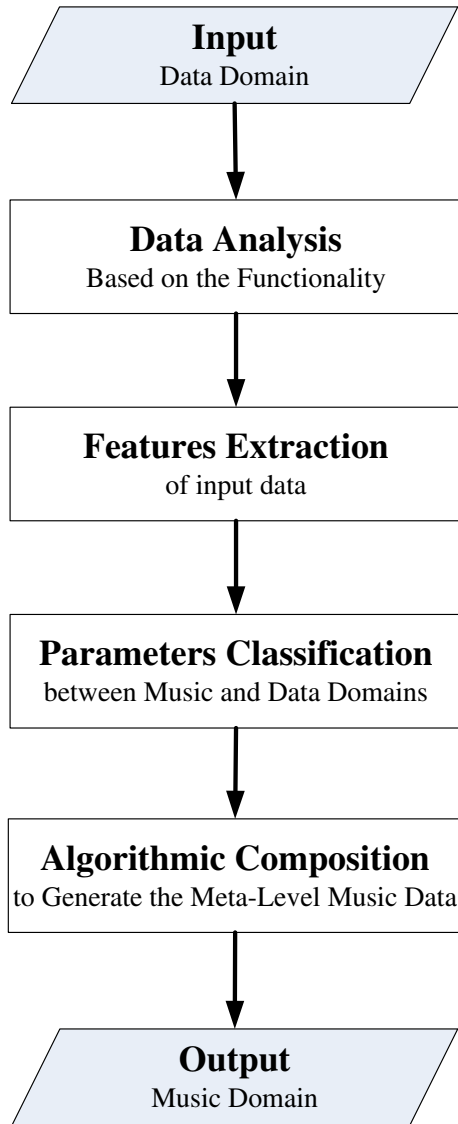
Throughout this section, a mechanism of mapping poetry data onto appropriate sound features is provided, followed by an overview of the system architecture.

3.1.1 Mapping recipe of Tx2Ms

The four tones in mid-ancient Chinese pronunciation are dichotomized into the only two categories in the classical Chinese intonation with the following characteristics, as shown in Table 2.

In addition to the intonation of the prosody, there are more features which could be taken into consideration, such as the semantic mood or style of the poetry (see Table 3), where “M” stands for mapping poetry parameter to musical parameter, and “n/a” stands for not available.

Fig. 5 General paradigm of sonification using algorithmic composition



For instance, the parameter of intonation (i.e., Ping or Ze) in the text domain is reasonably classified and mapped to multiple parameters in the music domain, where interval size refers to horizontal adjacent pitches and sonority refers vertical simultaneous

Table 2 Features of classical Chinese intonation system

Two categories	Four tones	Characteristics
Ping (level tone)	Level tone	Long, without any inflection
Ze (deflected tones)	Rising tone	Moving up
	Falling tone	Moving down
	Entering tone	Short

Table 3 Parameter classification and mapping of classical Chinese poetry

Parameter classification between two domains		Music					
		Rhythm	Interval Size	Sonority	Dynamics	Tempo	Mode
Poetry	Prosody intonation	M	M	M	M	n/a	M
	Poetic mood	n/a	n/a	n/a	M	M	n/a

itches, respectively. Table 4 shows the detailed mapping between poetic attributes and music parameters. The characters “C”, “J”, and “B” represent Chinese Pentatonic Mode, Japanese Hirajoshi Five-Tone Mode, and Balinese Gamelan Five-Tone Pelog Mode respectively. For rhythm, “sparse” means longer duration tone, and “dense” means shorter duration tone. For interval size, “small” means less than 3 semitones, and “large” means greater than or equal to 3 semitones. For dynamics, “soft” and “loud” represents the different sound volume levels.

3.1.2 Preliminaries of Tx2Ms

a. Use Markov chain in transition table construction

There are several divisions of techniques in algorithmic composition, inclusive of stochastic, rule-based flow control, grammar, chaotic and artificial intelligence. Markov chain is one of the stochastic processes in probability theory. In this work, Markov chain is the most convenient way to describe the intonation relations inside a poem. If a rule-based or a grammar controller is used, it will require more efforts to analyze the states of intonation transitions. If a stochastic or a chaotic mechanism is selected, then it may make choices without previous conditions, resulting in eventual uncertainty.

Accordingly, the following shows the Markov property, also known as Markovian.

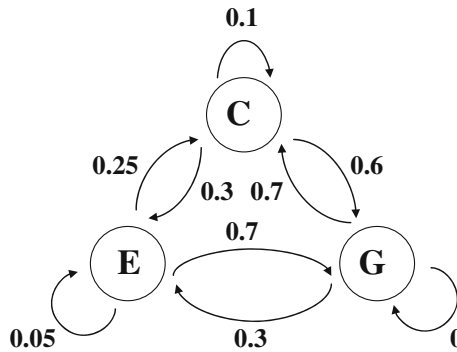
$$\Pr(S_{n+1} | S_n, S_{n-1}, \dots, S_1, S_0) = \Pr(S_{n+1} | S_n) \tag{1}$$

Further, the 1st-order Markov chain is described as a sequence of random variables $S_1, S_2, S_3, \dots, S_n$, with Markov process, where each S_i is one of the possible values from a state space S .

Table 4 Detailed mapping from poetry information to musical parameter

Parameter mapping		Music					
		Rhythm	Interval size	Sonority	Dynamics	Tempo	Mode
Prosody intonation	Ping	Sparse	Small	Harmonic	Soft	n/a	
	Ze	Dense	Large	Inharmonic	Loud		
	Tone	n/a					C/J/B
Poetic mood	Brightness	n/a			Loud	Fast	n/a
	Darkness				Soft	Slow	
	Neutrality or exoticism				n/a	n/a	

Fig. 6 Directed graph



Take two lower level musical elements for example. The state space of Pitch Class is {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11} while the state space of Rhythm could be {1/1, 1/2, 1/4, 1/8, 1/16, 1/32}. The following illustrates the process with Markov Property in Markov Chain.

$$S(t_0) \rightarrow S(t_1) \rightarrow S(t_2) \rightarrow \dots \rightarrow S(t_n) \rightarrow S(t_{n+1}) \tag{2}$$

A Markov chain could be represented either by a directed graph or a transition matrix. A directed graph consists of a set of states and a set of transitions with associated probabilities. A transition matrix of an N+1-dimensional probability table represents an N-state Markov chain, which tells us the likelihood of an event’s occurrence, given the previous N states [36, 51]. Figure 6 and Table 5 show the 1st-order Markov chain in terms of directed graph as well as Transition Matrix where C, E and G refer to the name of the tone (Do, Mi and So), respectively.

The transition matrix (or the stochastic matrix) *P* is the transition probability distribution, with *(i,j)*’th element of *P* equals to

$$p_{ij} = \Pr(S_{n+1} = j | S_n = i) \tag{3}$$

A vast majority of the uses of Markov chain in the algorithmic composition is to analyze and model the existing compositions. For example, some researchers have already analyzed the improvisation and chord progression by means of Markov model [8, 15, 17]. The Markov models used in these studies are mainly regarded as an analyzer or a model. In this study, we suppose the Markov model contributes to the stochastic algorithmic composition (see Fig. 7). The function of the Markov model is to facilitate meaningful mapping between poesy data and musical elements. With the advent of highly-relevant music, the emotional perception could be greatly improved during classical Chinese poetry appreciation. We extract the poetry styles and apply them into music features using stochastic algorithmic composition with the Markov chain.

In the preprocessing phase I of Tx2Ms, the interchanging Ping/Ze in each phrase is decomposed and aggregated for further recomposing of the rhythm sequence of original

Table 5 Transition matrix

Next Current	C	E	G
C	0.1	0.3	0.6
E	0.25	0.05	0.7
G	0.7	0.3	0

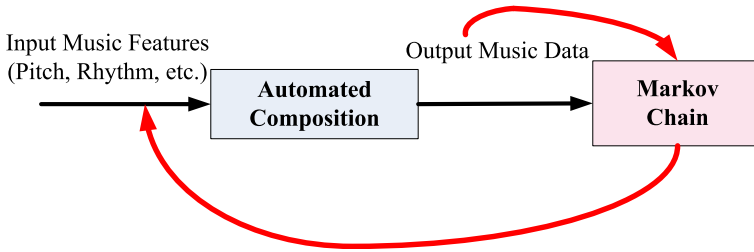


Fig. 7 The Markov model is developed for stochastic algorithmic composition

complete poem. Based on the rhythm sequence, a rhythmic transition table of the 1st-order Markov chain is computed for further algorithmic composition design. Figure 8 takes the famous Five-Character Quatrain, “Love Seed”, for example.

b. Apply Sieve Theory in pentatonic mode generation

Sieve theory in algorithmic composition is to define a set of diatonic musical scale by filtering out the unwanted notes from the chromatic scale. The concept of Pitch Class (PC) [37] is used for Sieve theory to classify the notes into a diatonic scale by “modulo 12” operation to obtain the Residue Class (RC) with integer numbers from 0 to 11.

The pentatonic or five note mode occurs in most of the ancient folk music in Asia. The prevalence of pentatonic modes in Chinese, Japanese, and Javanese music makes pentatonic modes have an Asian character for a long time. In particular, the pentatonic mode typifies the Chinese-style music since the traditional Chinese music is primitively based on pentatonic mode. Besides, a pentatonic mode, or, a five-tone mode, is a mode with five notes per octave. Table 6 and Fig. 9 present three typical pentatonic modes in the Tx2Ms (Table 7).

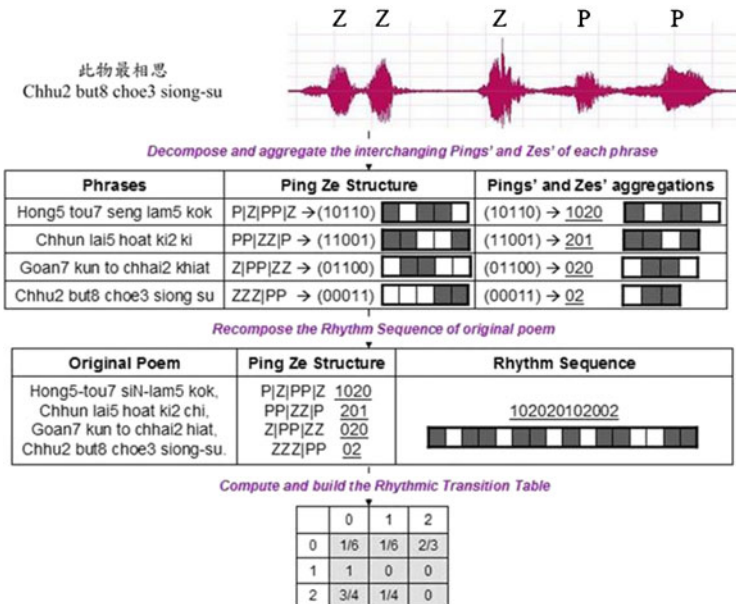


Fig. 8 An example of computing the rhythmic transition table from intonation

Table 6 Three pentatonic modes used in Tx2Ms

Pentatonic mode name	Pitch name	Pitch class (PC)
Chinese Pentatonic	(C, D, E, G, A)	PC _C : {0, 2, 4, 7, 9}
Japanese Hirajoshi Five-Tone	(C, D, E ^b , G, A ^b)	PC _J : {0, 2, 3, 7, 8}
Balinese Gamelan Five-Tone Pelog	(C, D ^b , E ^b , G, A ^b)	PC _B : {0, 1, 3, 7, 8}

Sieve theory is utilized here to generate pitches within a specific mode once the mode is recommended by *Tx2Ms*. Pitch class uses “modulo 12”. By using “mod 12”, any integer number above 12 should be reduced to a number from 0 to 11. This modulo operator can be visualized using a clock face (Fig. 10):

The function is described as below, where RP means Random Pitch (i.e., a random number integer) and $0 \leq RP \leq 127$; RC refers to Residue Class (i.e., the set of integers filtered), and is specified $RC = \{a, \dots, b\}$, where *a* is the minimum, and *b* is the maximum.

$$RC = RP \text{ mod } 12 \tag{4}$$

RC set of Chinese Pentatonic Mode (RC_C): $RP \text{ (mod } 12) = \{0, 2, 4, 7, 9\}$

RC set of Japanese Hirajoshi Mode (RC_J): $RP \text{ (mod } 12) = \{0, 2, 3, 7, 8\}$

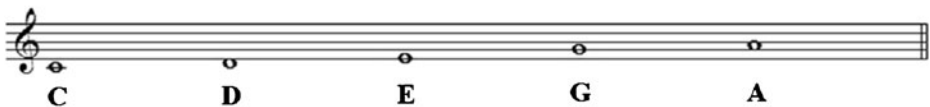
RC set of Balinese Pelog Mode (RC_B): $RP \text{ (mod } 12) = \{0, 1, 3, 7, 8\}$

For simplification, RC_C, RC_J, and RC_B are all named Pitch Class (PC) in ascending order of individual modes as PC_C, PC_J, and PC_B, respectively.

c. Speech-to-Mode Conversion based on Mapping Formants into Pitch Class

The formants of the top-5 significant and reliable phonetic segments are extracted to estimate the maximum likelihood of Pitch class among the three predefined pentatonic modes. Firstly, the two rhyming words and other three longest sounds are selected from all phonetic segments of the poem recitation. Secondly, formants in the vowels of the five words are analyzed with *Praat*, a free software for acoustic analysis (by Paul Boersma and David Weenink, Institute of Phonetic Sciences, University of Amsterdam). Then, each

Chinese Pentatonic Mode



Japanese Hirajoshi Mode



Balinese Pelog Mode



Fig. 9 The notation of the three typical pentatonic modes

Table 7 Significant phonetic segments extraction in “Love Seed”

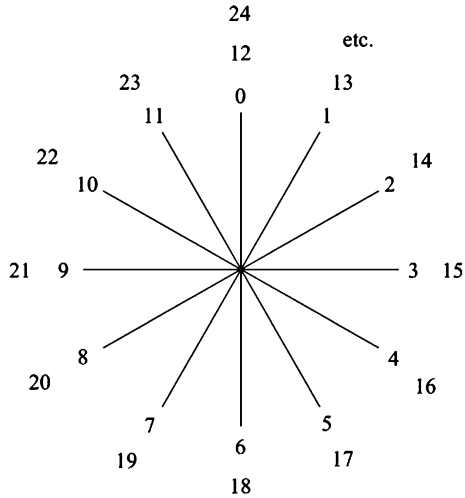
Phonetic Segments of the Poem “Love Seed” in Waveform¹ Display

<i>Text</i>	<i>Pronunciation²</i>	<i>Length (s)</i>	<i>Significance</i>
紅	hong5	0.32	
豆	tou7	0.47	long sound
生	seng	0.45	long sound
南	lam5	0.27	
國	kok	0.23	
春	chhun	0.27	
來	lai5	0.38	
發	hoat	0.28	
幾	ki2	0.25	
枝	ki	0.47	rhyming word
願	goan7	0.34	
君	kun	0.50	long sound
多	to	0.27	
采	chhai2	0.33	
擷	khiat	0.30	
此	chhu2	0.33	
物	but8	0.27	
最	choe3	0.36	
相	siong	0.42	
思	su	0.55	rhyming word

^a The digital speech file comes from National Digital Archives Program, Taiwan (recited by Tse-Nan Hung); URL: <http://dln.ntu.edu.tw/Education/94Web/7/index.html>

^b The pronunciation, most recited in “Southern Min” by the Roman Pinyin, derives from Feng-ju Lo: Tang Poetry Collection, URL: <http://cls.hs.yzu.edu.tw/300/ALL/ALLFRAME.htm>

Fig. 10 Pitch class of modulo 12



formant is converted into its approximate pitch based on the following equation to map a pitch’s fundamental frequency f (measured in hertz) to a real number p :

$$p = 69 + 12\log_2(f/440) \tag{5}$$

Afterwards, the Pitch Class of the real number p (further derived with modulo of 12) can be added by the Sieve theory:

$$PC = p \bmod 12 \tag{6}$$

The five derived pitch classes are arranged in ascending order. Thirdly, pitch class, $PC(p_1, p_2, \dots, p_n)$, is transformed to Interval Class, $IC(i_1, i_2, \dots, i_n)$, where

$$i_k = (p_{k+1} - p_k + 12) \bmod 12; \text{ if } k + 1 > n \text{ then } p_{k+1} = p_1. \tag{7}$$

The dissimilarities (distance) between the Interval Class of the poem and the interval class of the three predefined pentatonic modes are compared by Euclidean distance. The Euclidean distance between points $P = (p_1, p_2, \dots, p_n)$ and $Q = (q_1, q_2, \dots, q_n)$, in Euclidean n -space, is defined as:

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \tag{8}$$

Finally, the pitch class with minimum Euclidean distance is then selected as the best suitable mode for the particular poem. See the following two examples.

Example: “Love Seed” by Wang Wei (refer to Table 7)

- step1. Select the top-5 phonetic segments
- step2. Convert formants of vowels into pitch class
- step3. Calculate the interval class dissimilarity

The “Love Seed” pitch class is defined as $PC_X = (2, 2, 6, 7, 9)$. Transform Pitch Classes $PC_X(2, 2, 6, 7, 9)$, $PC_C(0, 2, 4, 7, 9)$, $PC_J(0, 2, 3, 7, 8)$ and $PC_B(0, 1, 3, 7, 8)$ to Interval Classes $IC_X(0, 4, 1, 2, 5)$, $IC_C(2, 2, 3, 2, 3)$, $IC_J(2, 1, 4, 1, 4)$ and $IC_B(1, 2, 4, 1, 4)$.

Compute the Euclidean Distance between $IC_X(0, 4, 1, 2, 5)$ and $IC_C(2, 2, 3, 2, 3)$, $IC_J(2, 1, 4, 1, 4)$, and $IC_B(1, 2, 4, 1, 4)$ as E_C , E_J , and E_B , respectively.

$$E_C = \sqrt{(0 - 2)^2 + (4 - 2)^2 + (1 - 3)^2 + (2 - 2)^2 + (5 - 3)^2} = 4$$

$$E_J = \sqrt{(0 - 2)^2 + (4 - 1)^2 + (1 - 4)^2 + (2 - 1)^2 + (5 - 4)^2} = 2\sqrt{6}$$

$$E_B = \sqrt{(0 - 1)^2 + (4 - 2)^2 + (1 - 4)^2 + (2 - 1)^2 + (5 - 4)^2} = 4$$

The similarity result is: Chinese Pentatonic \approx Balinese Pelog $>$ Japanese Hirajoshi. Thus, both PC_C and PC_B are recommended for the poem “Love Seed”.

d. Harmonic series utilization in sonority construction

Sonority is the result of two or more musical sounds combined simultaneously, different from melody, which comprises of two or more successive pitches as a temporal sequence. The harmonic series are utilized as a reference for sonority construction because even an untrained user can intuitively and conveniently detect by ear whether the sonority sounds harmonic or inharmonic. Taking the frequency of C1 (32.703 Hz) as an arbitrary fundamental frequency, frequencies of tones sharing the same relationships as harmonics in this harmonic series over the fundamental is calculated in a true overtone series as shown in Fig. 11. The indications added to individual notes imply each tone’s deviation in cents based on 1200 cents per octave standard.

The higher up the harmonic series, the more dissonant the sonority becomes. If the prosody intonation equals Ping then the sonority consists of intervals between lower successive harmonic series such as the P8, P5, and P4 (i.e., Perfect 8th, 5th, and 4th); if the prosody intonation equals Ze then the sonority consists of intervals between higher successive harmonic series such as the M3, and m3 (i.e., Major, and minor 3rd).

3.1.3 System architecture

Figure 12 illustrates the system flow chart of Tx2Ms, namely the text-to-music conversion of classical Chinese poetry, including raw data input, preprocessing, rhythmic Markov module, musification, meta-level score file, postprocessing, and audio file output.

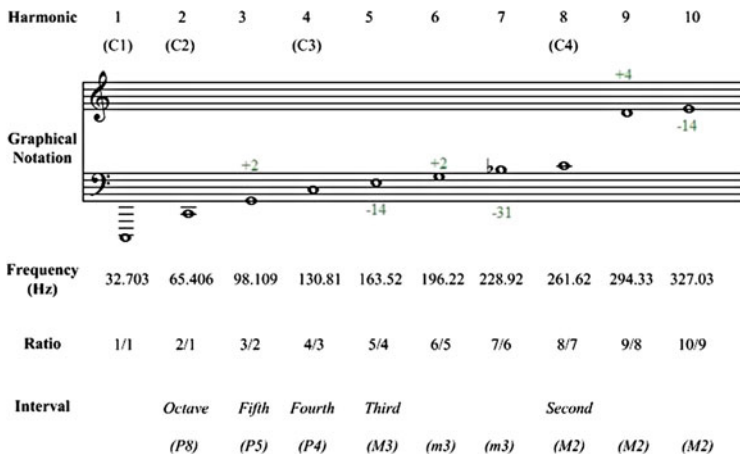


Fig. 11 The harmonic series over the fundamental frequency C1 (32.703 Hz)

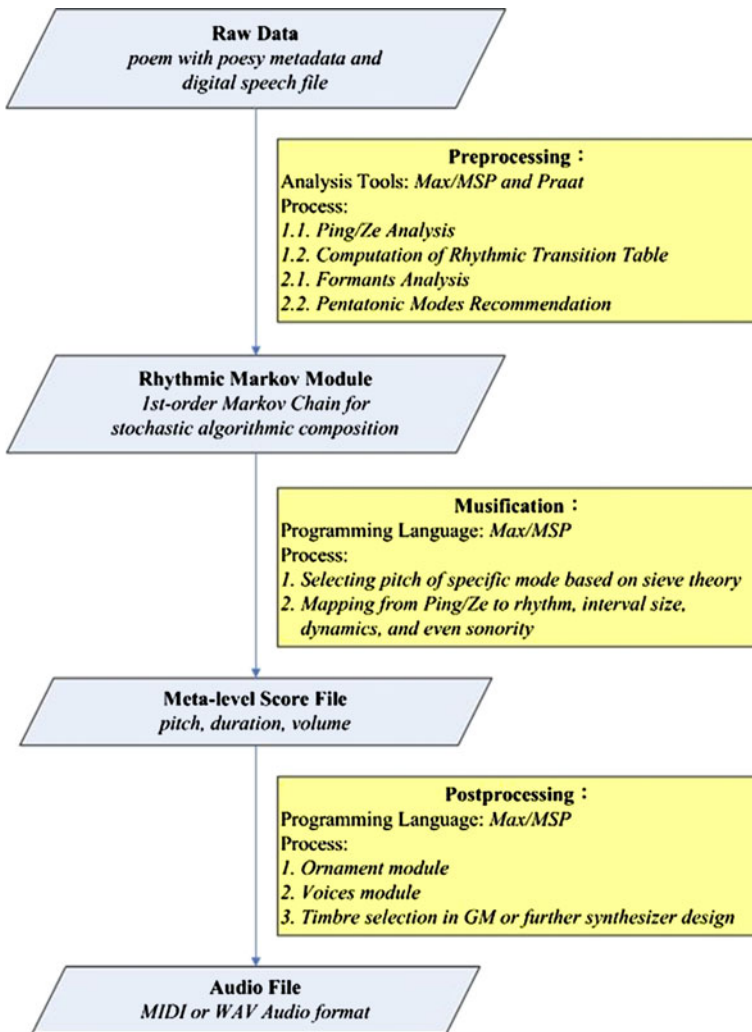


Fig. 12 System flow chart of Tx2Ms

Figure 13 shows the system architecture of Tx2Ms, including Preprocessing (Phase I), Musification (Phase II), and Postprocessing (Phase III).

In Phase II, the best suitable pentatonic mode is recommended according to the formants of prosody intonation. The mapping rules in the musification procedure are as shown in Fig. 14.

4 Experimental results and analysis

4.1 The implementation of the text-to-music mapping for classical Chinese poetry

The sonification scheme is implemented in Max/MSP for rapid prototyping and exploration, where Max/MSP is a graphical development environment for interactive computer music and multimedia, originally written by Miller Puckette and currently

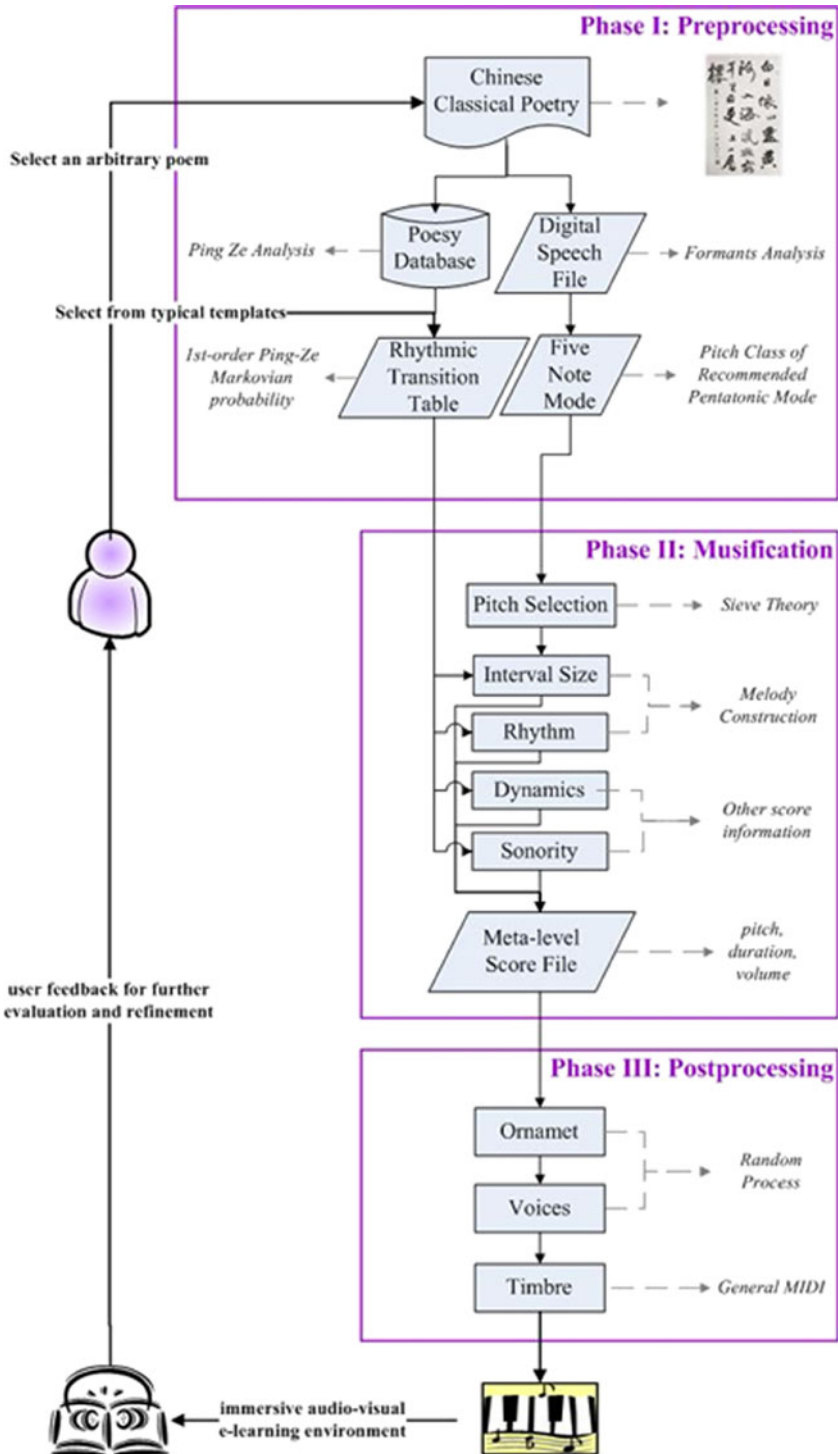


Fig. 13 System architecture of Tx2Ms

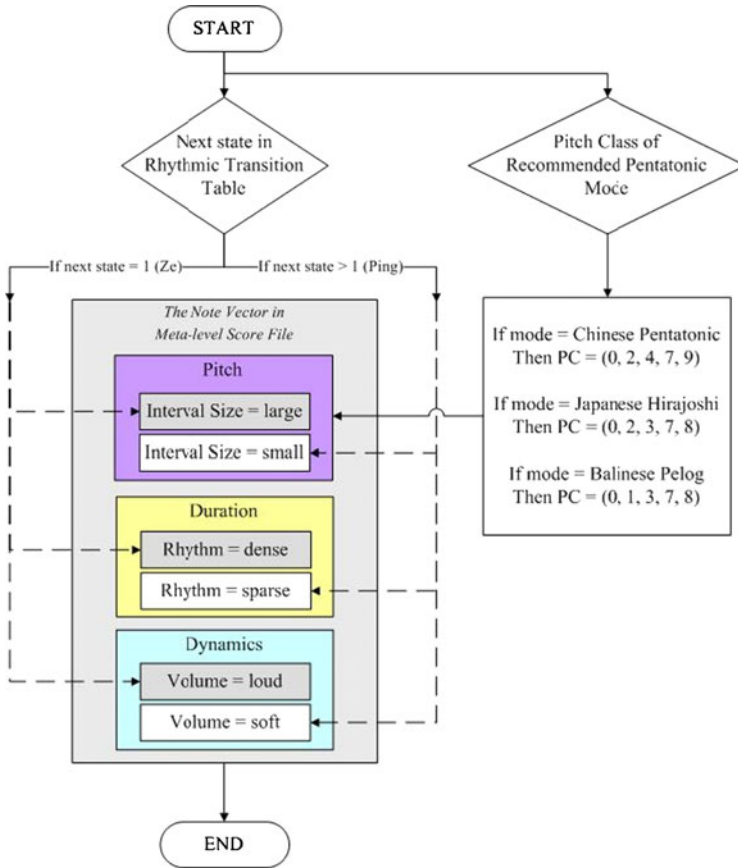


Fig. 14 The musification phase in Tx2Ms

developed and maintained by Cyling'74. The following figures depict the process of “Text-to-Music mapping of classical Chinese poetry”, Tx2Ms.

The implementation details of Tx2Ms are shown in Fig. 15, demonstrating different user-controlled level modules:

- **Mode Module:** To select the recommended or user-preferred Pentatonic Mode, with tonic and volume range selection. The default mode is Chinese Pentatonic mode.
- **Markovian Rhythm Module:** To select Ping-Ze based Rhythmic Transition table templates for specific poem. Although the default rhythm module is type I, this mode should be determined by the poetry type.
- **Rhythm-Controlled Interval Module:** The interval size is controlled by the transitional state result from the Markovian Rhythm Module. It does not require any input from users.
- **Sonority Module:** The sonority components are also controlled by the transitional state result from the Markovian Rhythm Module. It does not require any input from users.
- **Ornament Module:** User-controlled parameters for random process, including percentage of ornament, ornament duration, and ornament interval. All these parameters' range have been properly set, however, user can select them by their preferred styles.

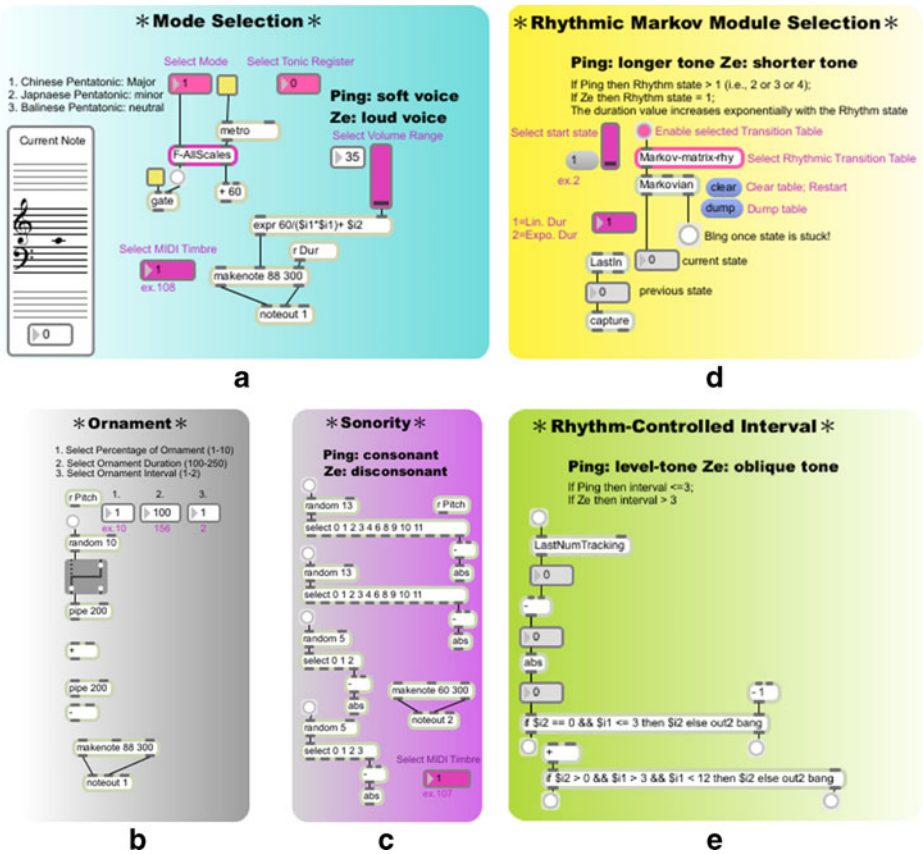


Fig. 15 Max/MSP implementation of Tx2Ms

Since Table 1 has illustrated the general Ping Ze for Five-Character Quatrain, Table 8 shows the rhythm sequence of each type of Five-Character Quatrain for building up the transition matrix.

Furthermore, Fig. 16 displays four basic templates for rhythmic matrices transferred from Ping Ze. Each message box has three numeric values, where the first and the second mean the current state and the next state, and the last value is the weighting of transition from current state to next state (Tables 9 and 10).

Figure 17 demonstrates a Ping-Ze score of part of the state sequence output (1, 3, 1, 3, 2, 1, 2, 1, 3, 1, 2, 1, 3, 1, 1, 3, 1, 3, 1, 3, 1, 3, 1, 1, 3, 2, 1, 3, 2, 1, 3) from Rhythmic Transition Table with a sixteenth note quantization, which is exemplified from the poem in Fig. 8.

Table 8 Significant formants in “Love Seed”

Text	Pronunciation	Formant f_1 (Hz)	Formant F_2 (Hz)	Formant F_3 (Hz)	MIDI note name	PC
豆	tou7	581	705	2955	D5 (587.3 Hz)	2
生	Sin	364	2101	2825	F#4 (370.0 Hz)	6
枝	Chi	437	2479	3196	A4 (440.0 Hz)	9
君	Kun	396	947	2777	G4 (392.0 Hz)	7
思	Su	295	2211	3139	D4 (293.7 Hz)	2

Fig. 16 Transition Matrix—four basic types of rhythm for the Five-Character-Quatrain in classical Chinese Poetry

Type I					Type III				
	0	1	2	3		0	1	2	3
0	0	2/5	0	3/5	0	0	1/3	0	2/3
1	1/2	0	0	1/2	1	1	0	0	0
2	0	0	0	0	2	0	0	0	0
3	1	0	0	0	3	1	0	0	0

Type II					Type IV				
	0	1	2	3		0	1	2	3
0	0	1/7	3/7	3/7	0	0	3/4	0	1/4
1	0	0	0	1	1	1	0	0	0
2	1	0	0	0	2	0	0	0	0
3	1	0	0	0	3	1	0	0	0

The following steps show one more experiment for the transformation process from Ping Ze to rhythm. Take Wang, Chih-Huan’s (王之涣) “Climbing the Stork Tower” (登鶴雀樓) for example.

- Step 1: Decompose phrases with Ping Ze Interchanging, where Z or 0 represents Ze, and P or 1 represents Ping.
- Step 2: Reconstruct the Rhythm Sequence, as shown in Fig. 18.
- Step 3: Build up the rhythmic transition matrix as shown in Fig. 19. based on the rhythm sequence acquired previously.

Table 9 Rhythm sequences of four Five-Character-Quatrain types

Type	Ping Ze ID	Ping Ze aggregation	Rhythm sequence
Type I	11001	201	20102020201
	00011	02	
	00110	020	
	11001	201	
Type II	11100	30	302020201
	00011	02	
	00110	020	
	11001	201	
Type III	00011	02	040402
	11001	201	
	11100	30	
	00011	02	
Type IV	00110	020	02020402
	11001	201	
	11100	30	
	00011	02	

Table 10 Ping Ze Interchanging for the Chinese Poetry “Climbing the Stork Tower”

Phrases	Ping Ze Structure	Ping Ze Interchanging → with Ping’s length
白日依山盡 pek8-jit8-i-san-chin7	仄仄平平仄 ZZIPPIZ	(010) → 020
黃河入海流 hong5-ho5-jip8-hai2-liu5	平平仄仄平平 PPIZZIP	(101) → 201
欲窮千里目 iok-kiong5-chian-li2-mok8	仄平平仄仄 ZIPPIZZ	(010) → 020
更上一層樓 keng2-siong7-it-cheng5-liu5	平仄仄平平 PIZZIPP	(101) → 102

4.2 Experimental evaluation and analysis

The system experimental evaluation is based on the system-generated music for classical Chinese poetry to see if it matches the imagery of the original Chinese poetry imagery. Figure 20 shows the experiment architecture with screenshots. The input data is the Chinese poetry, Five-Character Quatrain, “Love Seed”, then the system analyzes it to perform the sonification process based on the above mentioned method, and the output data is a sound file rendered from MIDI data.

All of the subjects are required to take 5-minute break as a baseline, without listening to the sonification wave file. A pair of earphones is used for the subjects to listen to the sound which has been rendered into wave files from MIDI format. General MIDI Instrument # 108 “Koto” is used for the proposed system. More detailed technical description about MIDI instrument definition and specification can be found in MIDI specification 1.0 [38].

A questionnaire survey was done to see if the generated music matches the corresponding Chinese poem imagery. Twelve people of the subjects have tried the Five-Character Quatrain, “Love Seed”, and the resulted automated composed music all seems to be acceptable. Likert 5-point scale (1: strongly disagree, 2: disagree 3: neither agree nor disagree, 4: agree, and 5: strongly agree) is used for the evaluation.

Subjects’ daily music listening time and music skill level are as shown in Figs. 21 and 22 respectively. It reveals that the subjects mostly have good musical talent, or are music lovers.

The item “As a whole the generated music matches the corresponding Chinese poem imaginary” is 3.75 points in the scale, and “the generated music emotion matches the mood of the corresponding Chinese poem” shows the highest score 3.90 in this questionnaire.

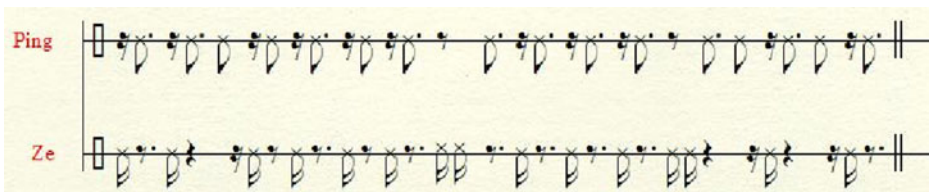


Fig. 17 Ping-Ze score of the poem “Love Seed” in Fig. 8

Fig. 18 Rhythm sequence of the Chinese poetry “Climbing the Stork Tower”

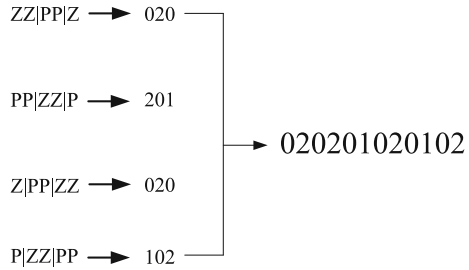


Fig. 19 The rhythm transition matrix of the Chinese poetry “Climbing the Stork Tower”

	0	1	2
0	0	1/3	2/3
1	1	0	0
2	1	0	0

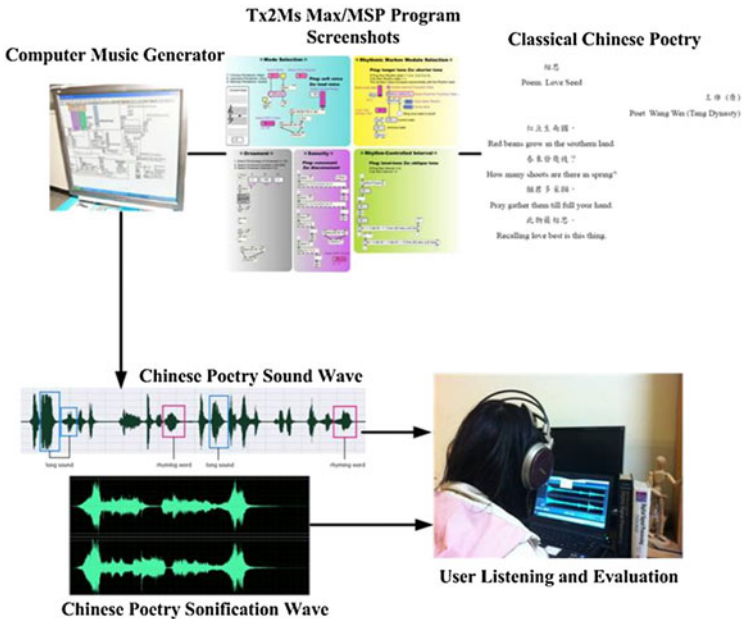


Fig. 20 System screenshots with user listening and evaluation

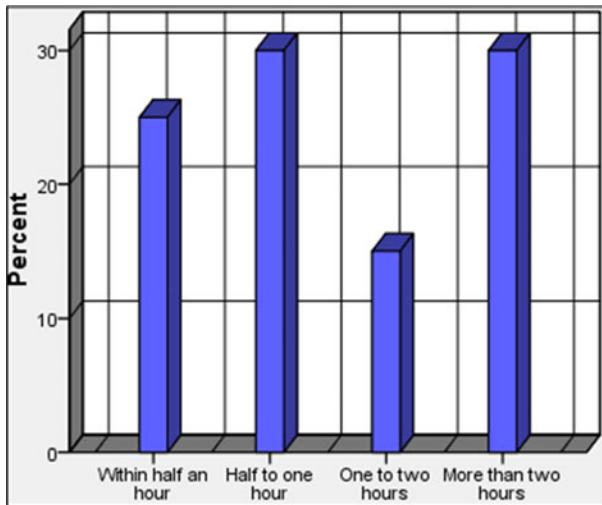


Fig. 21 Subjects’ daily music listening time

Reliability of the questionnaire is 0.712, which shows this scale with a good consistency. Please refer to Table 11 for more details.

In the aspect of “subjects’ daily music listening time”, the subjects who have longer daily listening time will give higher scores for the questionnaire than the subjects who have shorter daily listening time, as shown in Table 12.

A view from “subjects’ music skill level”, the subjects who have an expert-level skill will give higher scores for the questionnaire than the subjects whose skill level is normal. Years of musical training can be used to represent the level of musical skill, which can be found in much of the literature in the fields related to music psychology [19, 23]. Please refer to Table 13 in more details.

The result seems positive for most of both expert and non-expert listeners, although there are some music experts still criticizing the sound quality from algorithmic composition and computer generated music by the built-in Microsoft wavetable software

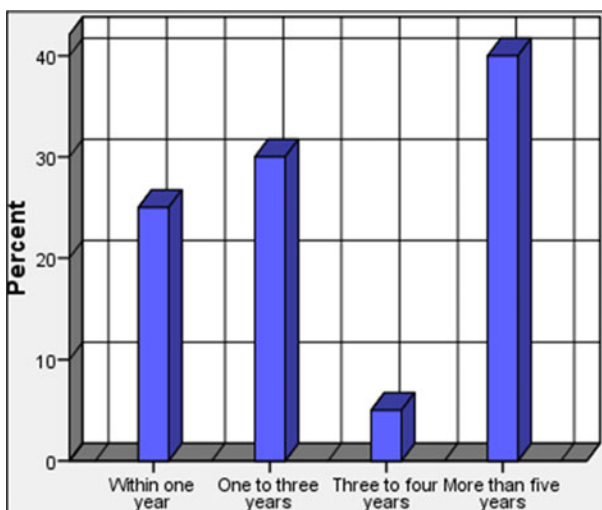


Fig. 22 Subjects’ music skill level

Table 11 Statistic correlation and reliability analysis for the generated music and classic Chinese poem

		The generated music rhythm matches the Ping-Ze rhyme of the corresponding Chinese poem	The generated music melody matches the pitch of the corresponding Chinese poem	The generated music emotion matches the mood of the corresponding Chinese poem	As a whole the generated music matches the corresponding Chinese poem imagery
N	Valid	20	20	20	20
	Missing	0	0	0	0
Mean		3.08	3.85	3.90	3.75
Mode		3 ^a	3	3 ^a	4
Std. Deviation		1.240	.933	.968	.851
Reliability Statistics (Cronbach's Alpha)		.712			

^a Multiple modes exist. The smallest value is shown

synthesizer. Even though most people already accept the computer music generated by software synthesizer via MIDI interface, there are still a few experts who cannot appreciate the generative sound by computer. Therefore the proposed system can be used for most listeners. Please refer to the following website to listen to the result, including the recitation sound of Chinese Classical Poetry—"Love Seed", Poet: Wang Wei (Tang Dynasty), and the Generated Music Based on "Love Seed": http://web2.cc.nctu.edu.tw/~eamusic/music_tech_lab/Chinese_Poetry/Chinese_Poetry.html

Based on our survey of the previous work, sonification can enhance the user's interest in the presented information in visual form. Multimodal senses will add synergies or create further informational dimensions. The difference between our work and most related researches includes the following: Firstly, previous work primarily converts graphical human-computer interface into non-speech sound; our work is to map characteristics of poetry into music. Secondly, most of the related researches treat sonification as the tool to generate audio signals; our work serves those who are not familiar with classical Chinese poetry with content and semantic context-based music which involves more music theory during the sonification process, instead of one-to-one mapping, such as color to pitch.

4.3 Discussion

The relationship between musical and prosodic structures has been developed by Lerdahl [32]; the poetry analysis can derive both rhythm and melody to demonstrate that a short poem can be entirely treated as music. Compared to the proposed system, this method is good for

Table 12 Cross-matching: daily music listening time v. s. questionnaire score

Subjects' daily music listening time		N	Mean
The generated music rhythm matches the Ping-Ze rhyme of the corresponding Chinese poem	Longer	9	4.33
	Shorter	11	3.45
The generated music melody matches the pitch of the corresponding Chinese poem	Longer	9	4.00
	Shorter	11	3.82
The generated music emotion matches the mood of the corresponding Chinese poem	Longer	9	4.22
	Shorter	11	3.36
As a whole the generated music matches the corresponding Chinese poem imagery	Longer	9	4.33
	Shorter	11	3.55

"Longer" means $\geq 1\sim 2$ h/day, and "Shorter" means $\leq 0.5\sim 1$ h/day.

"N" means the number of subjects.

Table 13 Cross-matching: music skill level v. s. questionnaire score

Subjects' music skill level		N	Mean
The generated music rhythm matches the Ping-Ze rhyme of the corresponding Chinese poem	Expert	9	4.33
	Normal	11	3.45
The generated music melody matches the pitch of the corresponding Chinese poem	Expert	9	4.44
	Normal	11	3.45
The generated music emotion matches the mood of the corresponding Chinese poem	Expert	9	4.33
	Normal	11	3.27
As a whole the generated music matches the corresponding Chinese poem imagery	Expert	9	4.56
	Normal	11	3.36

“Expert” means musical training time $\geq 3\sim 4$ years, and “Normal” means $\leq 1\sim 3$ years.

“N” means the number of subjects.

analysis, but its derived melody may be mostly atonal, and the “mapped” music is fixed. There are some tools to map data to auditory parameters and add context using the graphical interface. The AWESOME sound design tool has been developed for traffic situations using predefined harmonic complexity to generate sound [14]. Although it is good to use as an auditory signal, its result is not musical enough to perform the poetry sonification. The Sonification Sandbox [47] can generate MIDI output based on the settings of the Mappings Panel and the Context Panel. The Mappings Panel can control pitch, volume and pan; whereas the Context Panel controls a click track as notification using MIDI channel 10's percussive sound to generate tones. However the MIDI data is not music composition within a specific context. Most types of sonification theories and tools in the previous work are primarily used for general auditory display, rather than specific context application. The proposed Chinese poetry MIDI-based sonification is a specific customer designed tool, which can provide much more detailed context mapping information based on the poetry content, than most of the signal-based sonification methods.

5 Conclusion and future work

This paper introduces a method to generate music automatically based on the features of classical Chinese poetry: prosody intonation, poetic mood, waveform and frequency of pronunciation. Test results show that interesting and pleasant music segments can be produced in many cases. From a questionnaire, the generated music is considered to fit Chinese style. This composition approach has a high potential and is worth of further investigation.

Future work will continue to investigate the characteristics of other sonification applications, including e-book, Chinese calligraphy, multimedia animation, and computer game, in a style of creation for automated background music generation. As the results obtained in this stage are only music segments, it is desirable to integrate them into more complete pieces which can be used as the automated background music generation system. As both poetry and music are able to arouse emotional response, the application of emotional models is also under investigation.

Evaluation of the generated music is mainly via the students of the class at this point. A more systematic evaluation should be conducted to estimate the quality of generated music and more importantly, whether there is a strong association between the Chinese literature and its generated music.

The music generated by the system is according to the Ping-Ze rhyme and the tonality of the distinguished features from the Chinese poetic characteristics. For Text-to-Music conversion, not only the arrangement of text but also the pronunciation properties and the

syntactic characteristics of the poem are conveyed in the music output, based on the proposed mappings between the poetry and music.

Acknowledgement The authors would like to appreciate the support from National Science Council projects of Taiwan: NSC99-2410-H-155-035-MY2.

Appendix

The following is a translation of the poems exemplified in this paper, excerpted from the book, *Chinese-English Readings in poems of Tang dynasty and tsu of Sung dynasty* (中英對照讀唐詩宋詞), written by Shih, Ying-Chou (施穎洲).

相思

Poem: Love Seed

王維 (唐)

Poet: Wang Wei (Tang Dynasty)

紅豆生南國,

Red beans grow in the southern land,

春來發幾枝?

How many shoots are there in spring?

願君多采擷,

Pray gather them till full your hand.

此物最相思。

Recalling love best is this thing.

登鸛雀樓

Poem: Climbing the Crane Tower

王之渙 (唐)

Poet: Wang Chih-huan (Tang Dynasty)

白日依山盡,

The white sun sinks below the Mount.

黃河入海流。

The Yellow River flows to sea.

欲窮千里目,

To have full view of thousand li.

更上一層樓。

Climb up one story more must we.

References

1. Bain R (2006) The AIMS Project: creative experiments in musical sonification. International Computer Music Conference ICMC 2006, New Orleans, Louisiana, USA (November)
2. Ballas A (1994) Delivery of information through sound. In: Kramer G (ed) Auditory display: sonification, audification and auditory interfaces. Addison-Wesley, Reading, pp 79–94
3. Barrass S, Kramer G (1999) Using sonification. Multimedia systems 7(1) <http://www.springerlink.com/content/xd19ftjpf3pb30l/fulltext.pdf> (accessed 3 April 2011)
4. Blattner M, Papp L, Glinert P (1994) Sonic enhancement of two-dimensional graphics displays. In: Kramer G (ed) Auditory display: sonification, audification and auditory interfaces. Addison-Wesley, Reading, pp 447–470
5. Brewster A, Wright C, Edwards N (1993) An evaluation of earcons for use in auditory human-computer interfaces. In: Ashlund S, Mullet K, Henderson A, Hollnagel E, White T (eds) Proceedings of InterCHI'93. ACM, Amsterdam, pp 222–227
6. Chadabe J (1997) Electric sound: the past and promise of electronic. Music Prentice Hall, pp 178, ISBN 9780133032314
7. Chen H-H (2004) Introduction to composition in Chinese Poetry-Tsu. Wu-Nan Book Co., Taipei
8. Clement J (1998) Learning harmonic progression using Markov models. <http://www-lrn.cs.umass.edu/lab-lunch/papers/clement98learning.pdf> (accessed 10 May 2009)
9. Coagula software website: <http://hem.passagen.se/rasmuse/Coagula.htm> (accessed 8 December 2010)
10. Cope D (1992) Computer modeling of musical intelligence in experiments in musical intelligence. *Comp Music J* 16(2):69–83
11. Cope D (2004) virtual music: computer synthesis of musical style. MIT Press
12. Doornbusch P (2002) A brief survey of mapping in algorithmic composition. International Computer Music Conference, Göteborg
13. Durate J, Hsiao S-C, Huang C-F, and Winsor P (2006) The applications of Sieve Theory in Algorithmic Composition using MAX/MSP and BASIC. The 2nd International Conference WOCMAT, Workshop for Computer Music and Audio Technology, Taipei, Taiwan, pp 96–99 (March)
14. Fagerlönn J, Liljedahl M (2009) Awesome sound design tool: a web based utility that invites end users into the audio design process. In Proceedings of the 15th International Conference on Auditory Display (ICAD2009), Copenhagen, Denmark (May)
15. Farbood M, Schoner B (2001) Analysis and synthesis of palestrina-style counterpoint using Markov chains. International Computer Music Conference, Havana, Cuba, pp 18–22 (September)
16. Fowler A (1986) Franz Liszt's Petrarch Sonnets: the persistent poetic problem. *Indiana Theor Rev* 7(2):48–68
17. Franz M (1998) Markov chains as tools for jazz improvisation analysis. Master thesis, Virginia Polytechnic Institute and State University
18. Garzonis S, Jones S, Jay T, O'Neill E (2009) Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference. Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009, ACM, Boston, MA, USA, 4–9 April 2009
19. Gaser C, Schlaug G (2003) Brain structures differ between musicians and non-musicians. *J Neurosci* 23 (27):9240–9245 (October)
20. Goïna M, Polotti P (2008) Elementary gestalts for gesture sonification. Proceedings of the 2008 International Conference on New Interfaces for Musical Expression (NIME-08). Genova, Italy, pp 150–153 (June)
21. Hiller L, Isaacson L (1959) Experimental music. McGraw-Hill
22. Hsu C-Y (1997) Theory on classical Chinese poetry composition. Hungyeh Publishing, Taipei
23. Hunter P, Schellenberg E, Schimmack U (2008) Mixed affective responses to music with conflicting cues. *Cogn Emotion* 22(2):327–352
24. Hussein K, Tilevich E, Bukvic I, Kim S (2009) Sonification design guidelines to enhance program comprehension. In Proceedings of ICPC'2009, pp 120–129
25. Image-to-Sound Mapping website: <http://www.seeingwithsound.com/im2sound.htm> (accessed 8 December 2010)
26. Jeon M, Davison B, Nees M, Wilson J, Walker B (2009) Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies. In Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '09). ACM, New York, NY, USA, pp 91–98
27. Kaper G, Wiebel E, Tipei S (1999) Data Sonification and Sound Visualization. *Computing in Science and Engineering* 1, no. 4, <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=774840&isnumber=16814> (accessed 9 May 2009)

28. Kramer G (1994) Some organizing principles for representing data with sound. In: Kramer G (ed) Auditory display: sonification, audification and auditory interfaces. Addison-Wesley, Reading, pp 185–221
29. Kramer G (1994) An introduction to auditory display. In: Kramer G (ed) Auditory display: sonification, audification and auditory interfaces. Addison-Wesley, Reading, pp 1–77
30. Kramer G et al (1999) Sonification report: status of the field and research agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display. <http://www.icad.org/websiteV2.0/References/nsf.html> (accessed 3 April 2011)
31. Langston P (1989) Six techniques for algorithmic music composition. 15th International Computer Music Conference (ICMC) in Columbus, Ohio (November)
32. Lerdahl F (2001) The sounds of poetry viewed as music. *Ann N Y Acad Sci* 337–354 (June)
33. McAlpine K, Miranda E, Hoggar S (1999) Making music with algorithms: a case-study system. *Comp Music J* 23(2):19–30
34. McCormack J (1996) Grammar Based Music Composition. In: Stocker R et al (eds) *Complex systems 96: from local interactions to global phenomena*. ISO, Amsterdam, pp 320–336
35. Mitchell D (1985) Gustav Mahler: songs and symphonies of life and death: interpretations and annotations. University of California Press (November)
36. Moore R (1990) Chapter five: composing. In *Elements of computer music*. Prentice-Hall, Englewood Cliffs
37. Rahn J (1980) *Basic atonal theory*. New York: Longman; London and Toronto: Prentice Hall International, ISBN 0-02-873160-3. Reprinted 1987, New York: Schirmer Books; London: Collier Macmillan
38. Rothstein J (1995) *MIDI: A Comprehensive Introduction, Computer Music and Digital Audio Series, Vol.7, A-R Editions, Inc.*
39. Saue S (2000) A model for interaction in exploratory sonification displays. In *Proceedings of the International Conference on Auditory Display (ICAD2000)*, in Atlanta, Georgia (April)
40. Scaletti C (1994) Sound synthesis algorithms for auditory data representations. In: Kramer G (ed) *Auditory display: sonification, audification and auditory interfaces*. Addison-Wesley, Reading, pp 223–251
41. Shan K, Chiu S-C (2010) Algorithmic compositions based on discovered musical patterns. *J Multimed Tools Appl* 46:1–23
42. Spiliotopoulos D, Stavropoulou P, Kouroupetroglou G (2009) Acoustic rendering of data tables using earcons and prosody for document accessibility. In: Stephanidis C (ed) *Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction. Part III: Applications and Services (UAHCI '09)*. Springer, Berlin, pp 587–596
43. Sturm L (2001) Synthesis and algorithmic composition techniques derived from particle physics. *Proc. 8th Biennial Arts Tech. Symp.*, Connecticut College, New London
44. Vazquez-Alvarez Y, Oakley I, and Brewster A (2010) Urban sound gardens: supporting overlapping audio landmarks in exploratory environments. In *Proceedings of multimodal location based techniques for extreme navigation workshop, Pervasive, Helsinki, Finland*, pp 37–38
45. Verbeurgt K, Dinolfo M, and Fayer M (2004) Extracting patterns in music for composition via Markov chains. *Innovations in Applied Artificial Intelligence. 17th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE 2004 Ottawa, Canada*, pp 1123–1132 (May)
46. Verron C, Aramaki M, Kronland-Martinet R, Pallone G (2009) Analysis/synthesis and spatialization of noisy environmental sounds. In *Proceedings of the 15th International Conference on Auditory Display (ICAD2009)*, Copenhagen, Denmark, pp 36–40 (May)
47. Walker N, Cothran T (2003) Sonification sandbox: a graphical toolkit for auditory graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2003)*, Boston, USA (July)
48. Walker B, Kogan A (2009) Spearcon performance and preference for auditory menus on a mobile phone. In: Stephanidis C (ed) *Proceedings of the 5th International on Conference Universal Access in Human-Computer Interaction. Part II: Intelligent and Ubiquitous Interaction Environments (UAHCI '09)*. Springer, Berlin, pp 445–454
49. Winsor P (1989) *The computer composer's toolbox*. Windcrest Books, Blue Ridge Summit
50. Winsor P (1990) *Computer music in C*. Blue Ridge Summit, Penn. Windcrest Books

51. Winsor P (1992) Automated music composition. University of North Texas Press, Denton
52. Wu X, Li Z-N (2008) A study of image-based music composition. 2008 IEEE International Conference on Multimedia and Expo, ICME 2008 Proceedings, Burnaby, BC, pp 1345–1348
53. Yang C-C (2004) A study of constructing chinese poetry knowledge base—Su-Shi’s classical poetry knowledge base system. Master thesis, National Chiao Tung University



Chih-Fang Huang received both a Ph.D. in mechanical engineering and a master’s degree in music composition in 2001 and 2003 respectively from National Chiao Tung University. He is the Assistant Professor at the Department of Information Communication at Yuan Ze University, Taiwan, and the chairman of TCMA (Taiwanese Computer Music Association). His research papers include various fields, such as automated music composition and automatic control.



Hsiang-Pin Lu was born in Taiwan in 1966. He received the B.S. in computer science from the Chung Cheng Institute of Technology (CCIT), Taiwan, in 1988, and received the M.S. in computing from the Imperial College, London, in 1996. He is currently a Research Assistant at the National Chiao Tung University, Taiwan. From 1988 to 2006, and his current research interests including signal processing of FMCW sensor and algorithmic composition.



Jenny Ren received her B.S. in Computer Science and Information Engineering from National Central University in 2004, M.S. in Computer and Information Science and M.A. in Music both from National Chiao Tung University in 2006 and 2009, respectively. She is currently working for Multimedia Application Lab in Chunghwa Telecom Laboratories in Taiwan as an assistant researcher.