International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.11 (2015), pp.21-32
http://dx.doi.org/10.14257/ijmue.2015.10.11.03

Research on the Computer Music Production Technology System under the Digital Background

Yong Mei Li, Jingqiu Liu, Yi Wen Zhang, Chen Chen and Xuan Guan

Art institute, Hebei Normal University of Science and Technology, Qinhuangdao, 066004, China
2067377625@qq.com

Abstract

Computer music, as an important product of the information and digital age, belongs to the systems engineering combining the musical discipline with computer information control technology science. Through the applied research on digitalized acoustical signals, characteristic analysis of computer music, computer MIDI production technology, digital audio technology and computer music production software, this paper builds up a set of computer music production technology system based on a digital, informational and modernized background, and proves by experiments that this system makes the computer music production technology represent effects of simple and convenient usage, high-degree simulation and efficient extraction.

Keywords: *Computer; music production; MIDI; audio*

1. Introduction

Along with the rapid development of information technology, new media technology, especially computer music production technology develops continually and rapidly. Computer music, as an important expression of informational and digital age, also reflects to a great extent the social development and people's living standards. Especially with the constant appearance of new generations and the increased demand of aging populations for digital music, the requirements and renewal of computer music production technology become more and more urgent [1]. Computer music is an integrated discipline of systems engineering combining information science and music science. Particularly in the new media age, the various manners of communication as well as convenient and fast channels of communication can meet all kinds of personalized needs of modern people better. The existence of new media is based on technical support, and its rapid development is driven by digital technology, network technology and mobile technology. As the new media develop constantly, the digital audio constitutes an important part of new media information because of its convenient communication, various production technologies and means, high precision and high degree of controllability. Compared with the traditional analog audio signals, the digital audio signals have absolute advantages such as strong noise immunity, ease of storage and transmission, non-linear editing and post processing in bulk. Therefore, the digital audio production technology develops continuously in recent years, and its production means and methods change with each passing day. Computer music production technology includes not only the digital processing of audio, but also MIDI production technology [2-3]. Computer music production technology becomes the core technology of sound production, editing and processing in the new media environment under digital background. At present, all the sound contents spread through new media are achieved and spread quickly with the help of computer music production technology. We can say, the rapid development of digital

ISSN: 1975-0080 IJMUE

technology brings up new media, at the same time promotes the fast development of computer music production technology. Through the elaboration of digital and informational music and its production technology, this paper researches mainly on the music production system based on computer including the characteristic analysis of computer music, the digital audio technology, the computer music production software, and especially the computer MIDI production technology; builds up a set of computer music production technology system based on a digital, informational and modernized background; and based on the content-based music retrieval framework, constructs a music signature database by taking MIDI files as case study object to verify the effectiveness, immediacy and accuracy of this system [4-6].

2. Digital Music

Digital music refers to an art of music that is described, created, spread and stored via digital technology by using computer and Internet, and where sound streams are processed in other processing modes. It is a product of the electronic and computer age [7-8]. The appearance and development of digital technology is a revolution of musical art and sound processing technology since it changes the traditional manners of music creation, performance and production. As everyone knows, music is an art of sound, and the biggest difference between digital music and the traditional musical art lies in the manner of sound processing, producing and organizing. Meanwhile, as the digital music technology such as network develops, the music appreciation ways also change. Generally speaking, digital music is the music produced by computer via digital technology, whose production principle is as follows: use the digital synthesizer to phonate via computer's microprocessor; produce digitally described waveform in binary; modulate such digital waveform by using digital signal processor (DSP); transform digital audio signals into analog audio signals through digital analog converter (DAC) and send them to the amplifier for playing. Figure 1 is the schematic diagram of digital music composition.

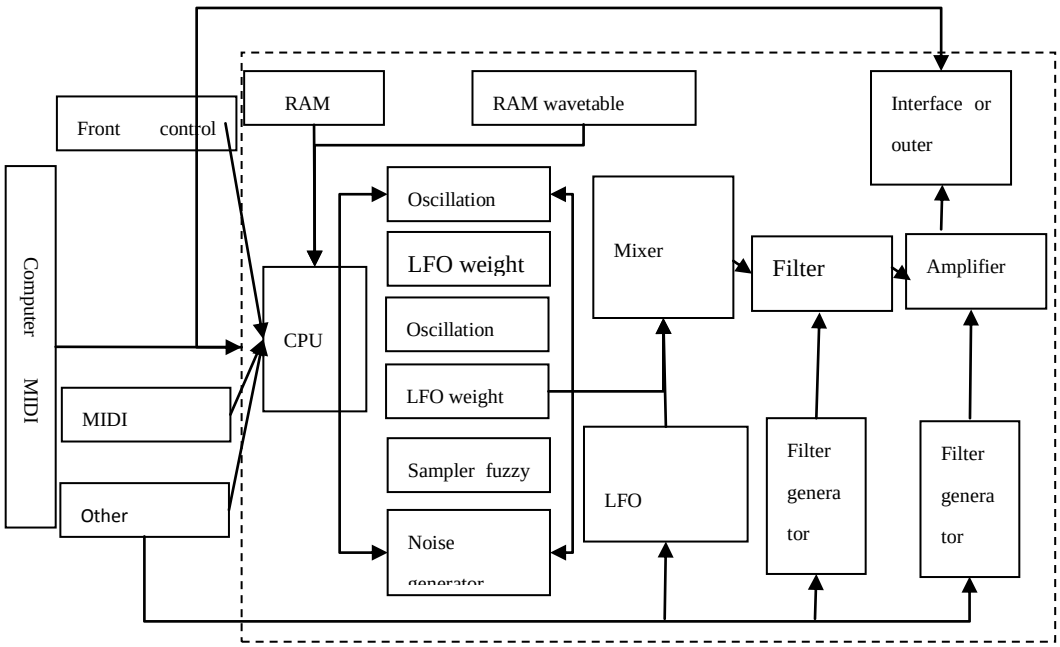

Figure 1. Schematic Diagram of Digital Music Composition

3. Computer Music Production Technology

3.1. Definition of Computer Music

Computer music is defined as the music that is produced with the involvement of computer [9]. Different from the traditional music production where analog signals are applied, the computer music production applies huge amounts of digital processing means which makes the forms and skills of music production more diversified and the quality more guaranteed. Nowadays, almost all the music information we hear is presented to us after computer's more or less processing. Currently, the audio information we can be in touch with in new media can be approximately divided into music, sound effect and commentary of audio clip. And the acoustic information must be digitalized by computer in order to be spread and applied by new media [10-11].

3.2. Contents of the Computer Music Production Technology

The computer music production technology mainly includes digital audio processing technology and MIDI production technology [12].

3.2.1. Digital Audio Technology

With the rapid development of computer technology, the quality of computer sound cards is further improved, and as a result, more sound-related development appears. TurboSynth, developing sounds from audio elements, was the originator of sound software with which users could create sounds they never heard. SoundTools became the first computer system that could authentically replace the two-track tape recorder, and it created a new world of digital editing, becoming a worldwide standard tool of mastering. Computer continued to accelerate at an amazing speed, and the sound card reached a standard specification--16-bit, 44.1 kHz sampling rate, digital I/0. Besides, the approach of combing audio track with MIDI sequencer was found out. By using the third-party high-quality audio interface, Studio Vision achieved the simultaneous running of audio track and MIDI track [13-14]. Then Digital Performer, Cakewalk4.0 on the Windows platform and some software products simply used for audio such as Deck and SAW were released. This type of software is easy to use. Figure 2 shows the Cakewalk audio equipment. Steinberg and Emagic added audio capabilities to the MIDI sequencer, and immediately Logic and Cubase became dominant. It was Digital design that originally promoted Pro Tools, a hybrid system which was established on the DSP basis. High-quality digitizer and DSP system removed lots of loads on computer, but ProTools was hard to be accepted by small recording studios as its price was exorbitant and it could be ran only on Macintosh platform. Steinberg promoted two standards of ASIO I/O and VST effect, and allowed other audio card manufacturers to create drives for its own audio cards in order to obtain the shortest latency time in Cubase environment. Figure 3 shows the Steinberg audio equipment. VST effect standard defines the CPU native effect processing standards. It is easy to achieve and apply, and inexpensive. While multi-track audio and MIDI software were competing to attract users, computer-based two-track editing and CD burning became the standard configuration of all recording studios, which was easy for all the typical computer systems. And the possibility of detailed editing of the software might be much bigger than that of any hardware equipment. In no time, the computer with high-quality audio card replaced the reel-to-reel tape and DAT. The VST and ASIO revolution brought some standards to the native system and changed absolutely the appearance of recording studios. Since then, many people abandoned completely the hardware synthesizer and recording equipment.

Figure 2. Cakewalk Audio Equipment

Figure 3. Steinberg Audio Equipment

3.2.2. MIDI Production Technology

MIDI is the abbreviation of Musical Instrument Digital Interface, which is used as a standard protocol of exchanging music information between the music synthesizer, the electronic musical instrument and the computer. Since early 1980s, MIDI has been gradually and widely accepted and applied by musicians, composers and musical instrument manufacturers [15]. MIDI, as a standard language used by electronic musical instruments and computers, is an agreement on messages (namely commands). Without generating signals, it transmits all kind of MIDI messages over MIDI cables, and the MIDI equipment or other electronic devices that accept MIDI messages will generate the sounds or execute the actions.

The MIDI interfaces on MIDI equipment generally include three different MIDI connectors: IN (input), OUT (output) and THRU (through connection). MIDI messages from MIDI equipment, sequencers and MIDI files are transmitted via the MIDI OUT [16]. Receiving equipment receives MIDI messages at MIDI IN port and then plays the sounds. MIDI source generates MIDI messages of binary data, and sends them to all synthesizers via MIDI message TURO box through MIDI cables. Synthesizers will generate the corresponding music in accordance with the MIDI messages, and send it to the speaker for playing. Figure 4 is the framework diagram of MIDI system.

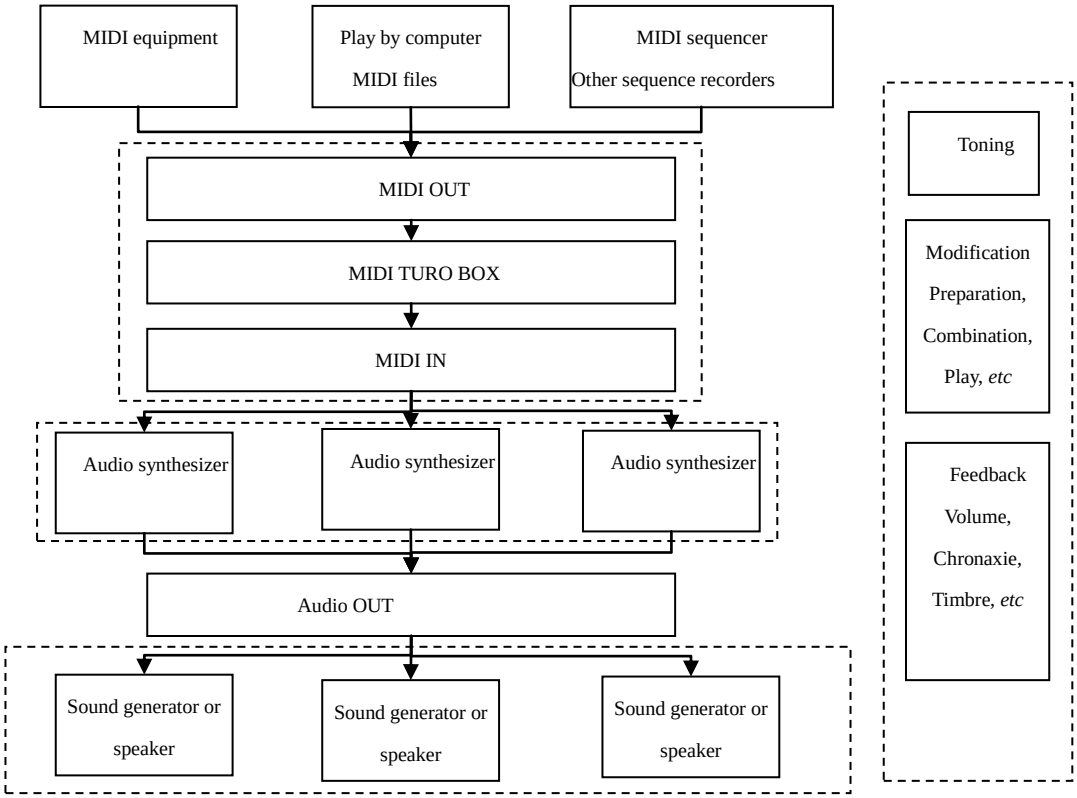

Figure 4. Structure and Process of the MIDI System Framework

4. Construction of the Computer Music Production Technology System

4.1. Hardware Configuration of the Computer System

At present, music production system is a typical hardware configuration with computer at the core for personal music production [17]. Its functions involve composition, arrangement, orchestration, recording, mixing, music printing and commercial music publishing, *etc*. And its peripherals usually include sound generator, synthesizer, M101 keyboard, sampler, digital or analog recording equipment, and monitoring equipment, *etc*. In accordance with the usage occasion and professional skills, the computer music system can be divided into the following three types.

Basic type: it mainly applies to entry-level computer music amateurs for non-commercial purpose. The system has lower requirements for computer: over 1G CPU GHz and 1GB memory [18]. It's enough for entry-level amateurs to learn and enjoy the primary MIDI production. As with the sound generator, the GM basic timbre solidified on computer sound card can be used in addition to the electronic keyboard with MIDI interface which is connected with the computer through MIDI cable. And with the keyboard of electronic keyboard, MIDI music can be simply produced on computer. Of course, the timbre quality of computer sound card and the keyboard's physical property of the electronic keyboard determine that music produced in such "workshop" is far worse than professional music.

Upgrading type: it is the most widespread and practical configuration mode at present in China. It includes such hardware as computer, synthesizer, sound generator, mixing console, monitoring equipment and digital recording equipment, *etc*. [19]. In this type, as the post analog acoustical signals need to be recorded and edited, it requires much higher computer hardware configuration: over 2.6G CPU, 2G memory as its lower limit,

high-capacity and better-performance hard disk to store massive audio data, and professional audio card as the interface (A/D, D/A conversion) of outer & inner computer audio data exchange. What's more, at least one or two professional sound generators or synthesizers are required, which are connected with the computer through special MIDI cable. With this set of system, professional MIDI music production as well as mixing recording, editing and producing of human voice and musical instrument can be achieved.

Professional type: this configuration is actually a miniature of computer-centered production system which is converted from the entire functions of traditional recording studio. In addition to the hardware referred in personal production type, its hardware configuration also include professional sound generator, sampler, sound workstation and digital recording equipment, *etc*. Sampler is kind of digital equipment used to obtain more species of timbres, more verisimilar sound effect and wider timbre controllability. In this system, the selection of audio hardware system is more important than that of MIDI system since the former directly influences the final sound effects of finished music. This system does not process MIDI information directly, but it has MIDI interfaces used for MIDI control function. Besides, it can be synchronous with other MIDI equipment, and it has functions of multi-channel simultaneous recording as well as high-performance sound analysis and processing. Generally it has sound workstations or external digital storage equipment around to store sound data files, which at the same time relieves computer's workload in the entire system.

4.2. Audio Production Software System

After the computer hardware is configured, the selection and application of software will seem to be especially important. The development of computer music software has experienced two phases. The first phase is the appearance of sequencer software based on computer and sequencer software, audio editing software and hardware such as sound generator and sound card. The second phase is a phase based on computer and sequencer software, soft sound generator and soft effect, and other plug-ins. The key of this phase lies in the appearance of "software package" which includes package of integrated sound generators and package of effects. HyperSonic2, as the most famous integrated sound generator software and with 2.0 as its latest version, is the most powerful wavetable sound generator software. Take Waves, the most professional wavetable sound generator software as an example, it is a complete set of effects package including 47 effects such as AUDIO TRACK 4-frequency band EQ (compression and threshold), C4 (compressor), L2 (limiter and dynamic processing), MAXXBASS (subwoofer processor), META FLANGER (stereo flanger), Q10 (10 band parameters EQ; compressor, equalizer, reverberator), ULTRAPITCH (6 band transposer), and NOISE (denoiser), *etc*. It's considered as a necessity of personal music studios. And among professional music instruments, it's the most popular software effect product as well as the most functional effects package. Figure 5 shows the HyperSonic2 software interface and Figure 6 shows the Waves C4 software interface.

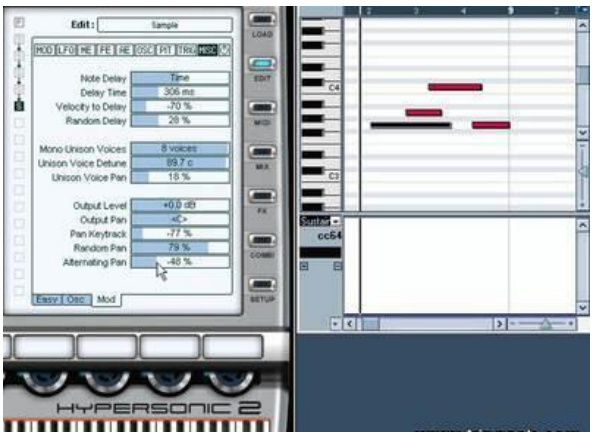

Figure 5. HyperSonic2 Software Interface

Figure 6. Waves C4 Software Interface

4.3. MIDI Music Database

MIDI music database is not only an important link but also a part with the highest use efficiency in building the computer music production technology system. The music database established by using MIDI extraction method can extract main melody eigenvectors from multi-track MIDI, set appropriate thresholds to conduct automatic music segmentation according to the duration distribution, and generate automatically mass music feature libraries. There are a good variety of MIDI file formats, and the following storage formats have been widely applied: RIFFMIDI (*.RMI) format released by Microsoft company, Standard MIDI FILE (*.SMF) format formulated by International MIDI Manufacturing Association, and some proprietary formats for MIDI music instruments such as the widely spread *.WRK format of Cakewalk. In PC machine, those files suffixed by .MID are all standard MIDI files. In general, standard MIDI file format can be widely supported. Except for special MIDI software, all the standard MIDI files can be opened, edited and saved in any MIDI software. MIDI files usually include multiple tracks with accompaniment, and it's very important to extract the main melody that represents the complete music information from the multi-track MIDI melody. Taking pitch as the key component of main melody eigenvector and as the melody information that describes music increases the rhythm information in the sound length's describing music. And a 2-dimensional ordered pair will be established.

Each note in melody corresponds to one feature point, which can be described as:

$$V =< [Pitch, time]$$

Thereof, Pitch refers to the value of pitch, and the note value varies from 0 to 127; Time is the improvement of MIDI tick, representing the information about sound length. The eigenvector corresponding to the note sequence of main melody can be expressed as:

$$V = \{v_1, v_2, v_3, v_4, \cdots, v_n\}$$

Thereof, V refers to the sequence of note feature points in the whole music, and n refers to the amount of notes.

Considering that there are phrases in music, the content features can be organized as per phrase, which will effectively help retrieve. The above vector can be further expressed as:

$$V = \{P_1, P_2, P_3, P_4, \cdots, P_k\}$$

Thereof, V refers to the sequence of note feature points in the whole music, and k refers to the amount of musical phrases.

$$P_i = \{v_{i1}, v_{i2}, v_{i3}, v_{i4}, \cdots, v_{in}\}$$

This eigenvector can express the melody and rhythm of music well.

Presently, most MIDI files downloaded from the network are multi-track files. Therefore, MIDI needs to be preprocessed in order to analyze the track features related to melody feature information from MIDI data. Based on the track features, build the corresponding linear weighting melody track extraction model; extract effectively the main track from multi-track MIDI files; remove the redundant information from multi-track MIDI; and extract a single track that represents the entire music information. Track features include mainly the following parts. Track name: generally there are such key words as “MELODIES”, “VOCAL” and “VOICE” in main melody name list; which in accompaniment list, there will be such key words as “ACCU”, “DRUM” and “BASS”. Channel number: there are 16 channels in MIDI; generally percussion uses the 10th channel as accompaniment, and melodic instruments use 1-9 and 11-16 channels. Balance degree of left & right sound channels x_1: the volume proportion pan value of left & right sound channels of each track is obtained via message “0xBn 0A value”; when the value is 64, the channel articulation of the track that main melody lies on is balanced. The calculation formula is shown in (1):

$$x_1(k) = pan(k) - 64/127, k = 1, 2, 3, 4, \cdots, 16 \quad (1)$$

The average key pressing strength x_2: the average key pressing strength of main melody is usually larger than that of other tracks. The calculation formula is shown in (2):

$$x_2(k) = \sum_{i=1}^{n} Vel(k, i) / N, k = 1, 2, 3, 4, \cdots, 16 \quad (2)$$

Thereof, $Vel(k, i)$ refers to the key pressing strength value of note i in channel k.

According to the track features, most MIDI main melodies can be determined. However, because of the complexity of MIDI files, sometimes the main melody cannot be correctly decided considering only the track features. Reference [13] points out that the algorithm of track selection depends to a large extent on the musical styles. If the main melody is observed in tracks, the selection of one main melody track will cause the loss of melody information. Selection of the optimal melody track based on hierarchical

clustering algorithm is to extract the main melody by selecting all tracks that contains melodies and to analyze and calculate the clustering distance of melody tracks by using mathematical model, which forms a process of main melody extraction method. And more meticulous work will be done in the process of determining the main melody. Based on Skyline Algorithm, this paper extracts the values of track and note with the highest pitch, and deletes other values of notes that articulate simultaneously. Through the analysis of MIDI contents, event sequence by time and pitch contained in the corresponding music fragments can be obtained. Details can be seen in Table 1: the corresponding MIDI data of music notes.

Table 1. The Corresponding MIDI Data of Music Notes

Event Num	Delte Time	Time Axis (Decimal)	Status	Pitch	Vel
1	0	0	9X (Note on in channel; note off in 0 key pressing strength)	64	5A
2	0	0		60	78
3	78	120		64	0
4	0	120		67	5A
5	3C	180		67	0
6	0	180		67	5A
7	3C	240		67	0
8	F0	240		60	00
9	0	240		67	5A
1	0	240		60	78
11	3C	300		67	0
12	3C	360		64	5A
13	3C	420		64	0
14	0	420		64	5A
15	3C	480		64	0
16	F0	480		60	0
17y	0	FF 2F 00			

According to Skyline Algorithm, those simultaneously articulating notes with smaller $pitch$ values can be removed and the MIDI event sequence will be obtained. Details can be seen in Table 2: the optimized results.

Table 2. The Optimized Results

Event Num	Delte Time	Time Axis (Decimal)	Status	Pitch	Vel
1	0	0	9X (Note on in channel; note off in 0 key pressing strength)	64	5A
2	78	120		64	0
3	0	120		67	5A
4	3C	180		67	0
5	0	180		67	5A
6	3C	240		67	0
7	0	240		67	5A
8	3C	300		67	0
9	0	360		64	5A
10	3C	420		64	0
11	0	420		64	5A
12	3C	480		64	0
13	0	FF 2F 00			

The above experimental data indicates that this algorithm uses MIDI analytical method to extract the melody eigenvector and to complete the automatic punctuation, after which the retrieval precision increases from 45.5% to 78.6% and the retrieval duration reduces from 6.331s to 4.321s. The greatly improved retrieval effectiveness facilitates faster and more effective system data structure, reduces tedious invalid notes, and establishes the data advantage for computer music production technology system.

4.4. System Pattern Analysis and Performance Results Evaluation

The IO operation of analog system may lead to such performance problems mainly about the import and export of files in ".mid" format which are influenced by the resource overhead of IO as well as the compilation overhead from files in ".mid" format to analog data structure. And the latter on can be improved: the analog data structure can be directly written out as Intermediary files through IO operation, and it will be compiled as files in ".mid" format only if necessary. However, compared with IO operation, the resource consumption of compilation process is still very limited.

From the performance constraints, it can be seen that the selection of containers has huge influences on the performance of the entire system. There are three containers for analog system: MIDI sequence container, MIDI event container and track container. Since the sequence lengths of different MIDI and the maximum quantity of tracks are knowable, these two types of containers can be achieved through simple expandable arrays. Therefore, the container of MIDI event becomes the key factor in influencing system performance.

Now this paper conducts a simple analysis of the following instance: assume that a user would create a sequence with one track that contains N events, and the absolute delay time for each event could be the same; after creation, the events would be processed with addition, deletion and modification at random positions. O(1) would be set as the time complexity of completing a simple operation (creation, modification, addition, deletion and event written-out). When implementing the scheme by using an array, L (L<N) would be set as the initial length of the array, and each expansion of the length would be 2 times

of the former. In creation, events need be sorted by time, and be expanded if necessary. By this time, the time complexity of creation process is O(NlogN). The time complexity is O(logN) in modification, deletion and addition processes when searching the target position by using bisection method. In comparison, these two data values are O(NlogN) and O(1) respectively when designing system with Hashtable where timestamps are used as keys. Thus it can be seen that a good selection of data structure has great influences on the system performance.

5. Conclusion

Computer music production system brings new concepts and new thinking methods to digital and informational music field, and it also brings new opportunities to the traditional vocal music production modes and means from the aspect of information science. Computer music production technology and multi-media technology not only can generate music and product audio products, but also can be widely used in music teaching, stage practices and other professional fields. For instance, the usage of software MIDISCAN can convert the music that is scanned to computer into practical sounds; the usage of recording software can convert audio files into music through recording method; the usage of accompaniment software can produce accompaniment disks needed in stage performance. In a word, the previously collective, multi-link and multi-discipline working process is simplified with high quality. This paper researched the digital processing of acoustical signals, hardware configuration in computer music system creation, software features, MIDI main melody extraction calculation method, sound encoder, tone generator module and especially MIDI database construction. It can be seen that the computer music production technology represents effects of simple and convenient usage, high-degree simulation and efficient extraction. And in this way, the music production technology system with high quality music under new media environment can also be produced. I believe that, through the continuous research and extension of computer music production technology system, the music production technology will develop more intelligently and human friendly, and will save more time, efforts and money for music production in the near future.

References

[1] Y. Jun, “Detail to MIDI Message and Standard MIDI-File Format”, Journal of South-Central University for Nationalities (Nat. Sci. Edition), Sup. Sep, vol. 22, **(2003)**.
[2] J. Rothstein, “MIDI: a comprehensive introduction”, Computer Music and Digital Audio Series, vol. 7, **(1995)**, pp. 23-27.
[3] H. Rothman, W. S. Brown, C. M. Sapienza and R. J. Morris, “Acoustic analyses of trained singers perceptually identified from speaking samples”, The Journal of Voice, vol. 15, no. 1, **(2001)**, pp. 25-35.
[4] R. Morales, “An Interactive Music Composition System Using Body Movements”, Computer Music Journal, **(2001)**.
[5] J. Fridrich “Miroslav Goljian Practical steganalysis of Digital Image”, State of Art SPIE, vol. 4675, no. 01, **(2002)**, pp. 1-13.
[6] D. Inoue and T. Matsumoto, “A standard MIDI files steganography and its power”, Transactions of IPSJ, vol. 43, no. 2, **(2002)**, pp. 2489-2501.
[7] D. Inoue, M. Suzuki and T. Matsumoto, “Detection resistant Steganography for standard MIDI files”, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E86A, no. 8, **(2003)**, pp. 2099-2106.
[8] X. P. Zhang and S. Z. Wang, “Dynamical running coding in Digital steganography”, IEEE Signal Processing Letters, vol. 13, no. 3, **(2006)**, pp. 1651-68.
[9] K. Yamamoto and M. Iwakiri, “A Steganography to music Code with adaptation in musical expression”, Transactions of Information Processing Society of Japan, no. 478, **(2006)**, pp. 2724-2732.

[10] M. Iwakiri, K. Yamamoto and K. Sekine, "Watermarking Techniques to Computer Music Codes Al. Half scrambling", Transactions of Information Processing Society of Japan, vol. 43, no. 2, **(2002)**, pp. 225-233.
[11] A. R. Burton and T. Vladimirova, "Generation of musical sequences with genetic techniques", Computer Music Journal, vol. 23, no. 4, **(1999)**, pp. 59-73.
[12] M. Bigerelle and A. Iost, "Fractal dimension and classification of music", Chaos, Solitons and Fractals, vol. 11, no. 14, **(2000)**, pp. 2179-2192.
[13] M. Kermit and J. Eide, "Audio signal identification via pattern capture and template matching", Pattern Recognition Letters, vol. 21, no. 3, **(2000)**, pp. 269-275.
[14] L. Yang, W. Y. Dong and L. B. Long, "A new method for approximate melody matching and Its application in query by humming system", Journal of Computer Research and Development, (in Chinese), vol. 40, no. 11, **(2003)**, pp. 1554-1560.
[15] Hirata K. and Matsuda S., "Interactive music summarization based on generative theory of tonal music", Journal of New Music Research, vol. 32, no. 2, **(2003)**, pp. 165-177.
[16] Raphael C., "A probabilistic expert system for automatic musical accompaniment", Journal of Computational and Graphical Statistics, vol. 10, no. 3, **(2001)**, pp. 487-512.
[17] Goto M. and Muraoka Y., "Real time beat tracking for drumless audio signals: Chord change detection for musical decisions", Speech Communication, vol. 27, no. 3/4, **(1999)**, pp. 311-335.
[18] Cope D., "Recombinant music: using the computer to explore musical style", Computer, vol. 24, no. 7, **(1991)**, pp. 22-28.
[19] Lu L., Zhang H. J. and Li S. Z., "Content based audio classification and segmentation by using support vector machines", Multimedia Systems, vol. 8, no. 6, **(2003)**, pp. 482-492.

