Keynote-Dependent HMM Based Musical Chord Recognition Method

Da-chuan Wei
Science and Technology Industry Division,
Jilin Architectural and Civil Engineering Institute,
Changchun, China
Email: weidachuan2010@163.com

***Abstract*—Chord sequences and chord boundary is the mid-level performance of the music signal compactness and robustness. Automatic chord recognition is very attractive to researchers in the field of music information retrieval. To improve accuracy of musical chord recognition algorithm, in this study, the importance of keynote was fully considered. According to the music theory, 24 keynotes were defined. For each keynote, a Hidden Markov model was established, which is called keynote-dependent HMM. And then, a recognition algorithm of music chord based on keynote-dependent HMM was proposed. In this algorithm, use MIDI music corpus to train the keynote-dependent HMM, and to improve the recognition rate and facilitate the calculation, a 6-dimensional vector of tonal centroid is used as the feature vector. The experimental results showed that the proposed keynote-dependent HMM had better recognition effect than that of keynote-independent model.**

***Index Terms*—musical chord recognition, keynote-dependent HMM, MIDI, tonal centroid**

I. INTRODUCTION

Early chord recognition system from the perspective of the natural raw audio.Some systems, the application of a simple sample matching algorithm, which is built on pre-defined chords templates.The other system is more complex machine learning techniques, such as hidden Markov models (HMMs) or support vector machine (SVMs).

The great success of the hidden Markov model (HMM) in speech recognition, allows researchers to start thinking about the hidden Markov model is applied to a musical chord recognition. Have a lot of chord recognition system began to create a simple hidden Markov model, but it did not take into account the impact of the music tone for chord recognition. Tone in music is a very important property, this article give full consideration to the importance of tone, musical chord recognition system to be achieved by the introduction of the tone information to build the key-dependent HMM-based musical chord recognition system.

Chord is one basic element of the western tempered music system, which has dominated the mainstream music for centuries. It is made up of several notes sounding at the same time, and the changes in the melody are reflected by the balanced degree and the discretion differences between notes [1]. With the development of information technology, how to make computer accurately "understand" music content in human and make the communication between human and computer as easy as that between people, has been one goal people are pursuing. Therefore, music chord recognition technology has become a new research topic.

As the basis and hot topic of music signal processing by computer, the research on automatic chord recognition was into a climax in the late 1990s both at home and abroad. Su and Jeng proposed a chord classification method using wavelet transform and self-organized map neural networks [2]; Cabral introduced a chord recognition algorithm based on audio descriptor — Extractor Discovery System(EDS) [3]; Takuya Yoshioka presented an automatic chord transcription with concurrent recognition of chord symbols and boundaries, whose core is the assume search algorithm [4]. Sheh and Ellis used statistical method to recognize and segment the chords, which trained the Hidden Markov Model (HMM) by the EM Algorithm [5]; Bello and Pickens conducted a similar study, they not only avoided the randomness of the initialization but also updated the parameters optionally, which improved the recognition rate [6]; Gomez proposed Harmonic PCP(HPCP) feature to recognize chords, and the recognition system with the new features achieved good effect [7]; Li used wavelet packet analysis technology to do three-string wavelet packet decomposition for chords played by guitar, and combining with the envelope extraction technology based on Hilbert transform, gained the contour time-frequency plan of the time-domain reconstructed signal of the wavelet coefficients of the chords [8]; Sun presented a real-time perception method for chords based on artificial neural network (ANN), according to the principles of psychology for music cognition [9].

In Western music, the keynote and chord are the two very important properties. The music keynote defines the reference point and the tone center of the music, which can be used to handle the musical tunes, harmony, and rhythm and so on. To improve accuracy of musical chord recognition algorithm, in this study, the importance of keynote was fully considered, and a recognition algorithm of music chord based on keynote-dependent

doi:10.4304/jsw.8.4.916-923

HMM was proposed. Firstly, by harmonic analysis, the marked chords can be gained from the MIDI music files, and then the feature vectors can be extracted according to the audio synthesized on the above file; Secondly, according to the music theory 24 keynotes were defined, and for each a Hidden Markov Model (HMM) [10] was established, called keynote-dependent HMM, and then use the extracted feature vectors and script files as input to train the model; Finally, when inputting the music audio, select a maximum possible tone model from the above 24 models by Viterbi decoding. This algorithm can not only determine the keynote of the input music, but also get the chord sequence from the optimal state path corresponding to the keynote model.

II. Algorithm Design

Chord recognition system studied in this paper relates to two kinds of feature vectors. The first characteristic vector referred to as chroma vector (chroma vector), most of the chord recognition system are the application of this feature vector or its deformation. The second feature vector is called the pitch from the heart (tonal centroid), referred to as self-aligning feature vector system to achieve this is to use this vector as a feature. Firstly, the chroma vectors to achieve the chord recognition system, and then the tone from the heart vector. Finally these two different feature sets were compared.

MIDI is short for Musical Instrument Digital Interface, intended for Musical Instrument Digital Interface. It is a unified exchange agreement between an electronic musical instruments and electronic instruments and computer. Through it, various MIDI devices can accurately transmit MIDI messages. Come out since the early 1980s, it experienced a long period of development in the broad sense we can interpret it as electronic music synthesizer, computer music collectively, including all the technical protocols, devices, etc..

MIDI file format is also known as a frequency synthesizer audio files, it is by the frequency synthesizer audio files stored code (control) to control the Sound Blaster audio synthesizer output audio signal encoding and time simply is itselfnot sound, the sound conversion by the A / D driver sound card.

It has the advantage: the MIDI file does not contain the information of the streaming media, file size is smaller, the sound quality is quite good. The disadvantage is that the audio file: MIDI files are not on the real meaning, but the audio control file does not support live original music or vocals.

A. Feature Extraction

Chroma Figure (chromagram,) was first proposed by Fujishima. It is defined pitch the concept of a standard method. It is a function of the audio and high school two attributes "height and chroma. Pitch height of the vertical movement of the octave, used to indicate the pitch belongs to which a tone in which an octave. Pitch chroma that the relationship between the location of the pitch and octave pitch.

Degrees in Figure B-dimensional vector, and represents a chroma. Where B is the number of octave scale, said the 12 semitones in the chromatic tension. Since the chord is composed by a number of tones and chords mark can only be determined by the pitch in the chromaticity location, so do not consider them the height of the chroma vector can be seen can be used to represent musical chords or music idealized characteristics of the tone.

Chroma vector is calculated as follows:

Firstly, the input signal $X(k)$ constant Q transform and get X_{cq}.

$$X_{cq}(k) = \sum_{n=0}^{N(k)-1} w(n,k)\,x(n)\,e^{-j2\pi f_k n} \tag{1}$$

where, the $w(k)$ is the analysis window, the length of $N(k)$.

The center frequency f_k is defined as follows:

$$f_k = 2^{k/\beta} f_{\min} \tag{2}$$

Where β is the constant Q transform an octave interval the number of alto.

Secondly, calculated X_{cq}, you can get the chroma vector of the Chroma.

$$Chroma(b) = \sum_{m=0}^{M-1} \left| X_{cq}(b + m\beta) \right| \tag{3}$$

Where b = 1, 2, 3 ... beta, M is the number of octave constant Q spectrum.

In our chord recognition algorithm, use tonal centroid to be as the feature vector, which a 6-dimensional feature vector proposed by Harte, can be used to detect the changes in harmony of music audio [11]. Tonal centroid vector is built on the basis of harmony network or audio network. Harmony network or audio network is the plane of the pitch relations, said. In this plane, the close relationship between the pitch, such as the five-octave or large/small octave, the Euclidean distance between them is very small.

In theory, harmony network is an infinite plane, that is, there are numerous pitch. Assume different character homonym and octaves are equivalent, that is only 12 kinds of pitch classes, then the infinite plane can be five-tone ring, replace the junior tone ring and a small three tone ring, thus established a three-dimensional toroidal. As shown in Figure1, the 6-dimensional tonal space can be as three tone rings made up of Fifths tone ring, Minor Thirds tone ring and Major Thirds tone ring. The numbers in the tone ring are corresponding to each type of pitch and pitches latest to it, and the tonal centroid of A's Major Thirds tone is at point A.

If the C sounds as pitch 0 reference point, then the five-octave ring there are 12 different points, 0-7-2-9-... -10-5, and finally loop back to 0. However, if the three tone ring mobile, requires only three steps back to the reference point, ie, 0-3-6-9-0. Junior tone ring is also defined in the same way. Thus, the 6-dimensional vector can be viewed as combination of (x_1, y_1), (x_2, y_2) and (x_3, y_3).

As mentioned above, all of the pitches in a chord can be viewed as a single point in the 6-dimensional space, the procedure of feature extraction is shown in Figure 2.

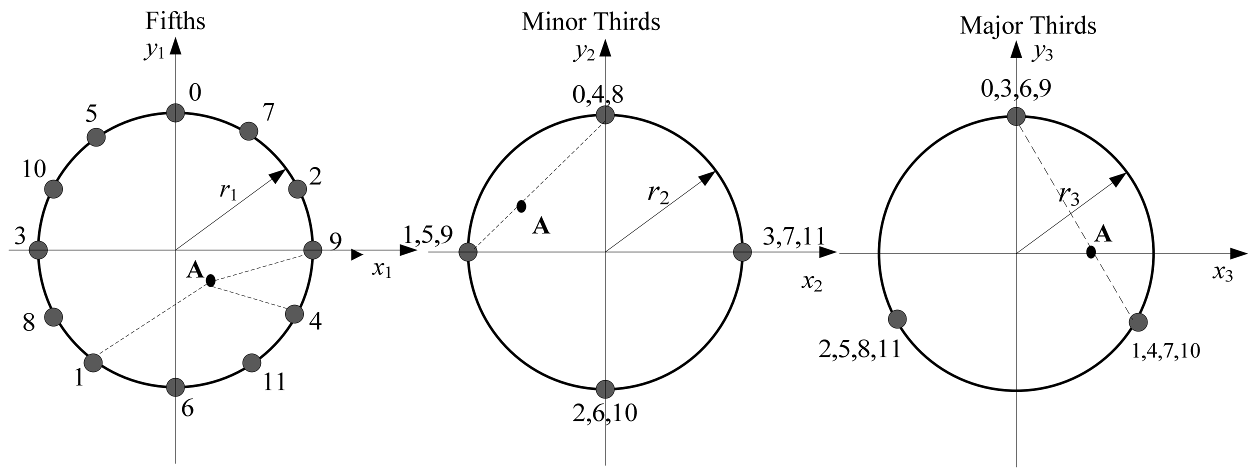

Figure 1. Schematic of three-dimensional toroidal

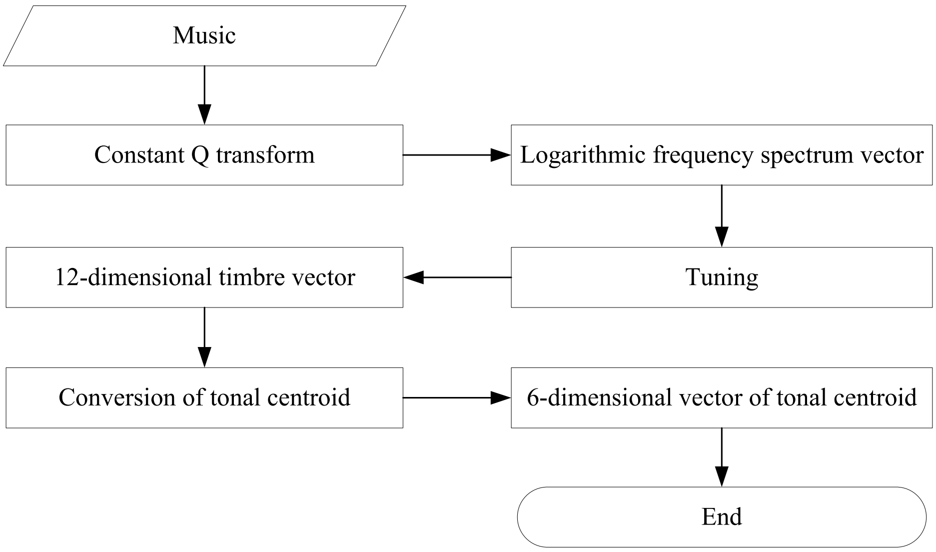

Figure 2. The extraction procedure of 6-dimensional tonal centroid vector

Harte and so on through the 12-dimensional chroma characteristics of projection will be an adjustment to the above audio network equivalent of the three rings, resulting in a six-dimensional tonal centroid vector. In other words, each chromatic pitch corresponds to a point on the three rings. In this study, the 6-dimensional tonal centroid vector can be calculated by multiplication between a transformation matrix with size of 6×12 and a timbre vector with size of 12×1, as shown as (1).

$$\zeta_n(d) = \frac{1}{\|c_n\|_1} \sum_{l=0}^{11} \Phi(d,l) c_n(l) \qquad \begin{array}{l} 0 \le d \le 5 \\ 0 \le l \le 11 \end{array} \tag{4}$$

Where $\Phi = [\varphi_0, \varphi_1 \cdots \varphi_{11}]$ is the transformation matrix, and:

$$\varphi_l = (\Phi(0,l), \Phi(1,l), \Phi(2,l), \Phi(3,l), \Phi(4,l), \Phi(5,l))^{\mathrm{T}} \tag{5}$$

$$\Phi(0,l) = r_1 \sin l \frac{7\pi}{6}$$

$$\Phi(2,l) = r_2 \sin l \frac{3\pi}{2}$$

$$\Phi(3,l) = r_2 \cos l \frac{3\pi}{2}$$

$$\Phi(4,l) = r_3 \sin l \frac{2\pi}{3}$$

$$\Phi(5,l) = r_3 \cos l \frac{2\pi}{3}$$

Where $0 \le l \le 11$, r_1, r_2 and r_3 are the radius of the three rings in Figure 1 respectively, and r_1=1, r_2=1, r_3=0.5.

For MIDI files synthetic music and audio, we use Timidity + + to get the job done. Timidity + + is a compositing software, which can convert MIDI files to WAVE form audio files.

For the MIDI file access, GUS (Gravis, Ultra Sound) voice synthesis. First, we used to synthesize classical music, piano, violin, viola and cello. Then, organ, electric guitar, electric bass and orchestra to the synthesis of the Beatles music.

Feature extraction, the frame length of 743ms, the frame shift is 185ms. In the training model, it is necessary to preprocess the training data set, and then the 12-dimensional chroma features to convert the six-dimensional tonal centroid features.

B. The Establishment of Keynote-dependent HMM

Chord recognition algorithms are just the type of music classification, hidden Markov model without taking into account the tone for chord recognition. Unlike the previous algorithms, in our algorithm firstly the Hidden Markov Model (HMM) is established, and then the keynote information is added, and on this basis, the keynote-dependent HMM is established.

① Establish HMM

As mentioned earlier, music is divided into two categories, namely classical music and Beatles music. Therefore, when conducting chord recognition, it is necessary to these two categories of music were the establishment of the HMM.

- Establish HMM with 36 states for the training data set;
- In the HMM, each state represents a single chord;
- The observation distribution of HMM obey a multi-value single Gaussian distribution, which is 12-dimensional for timbre feature and 6-dimensional for tonal centroid;
- This multi-value single Gaussian distribution is defined by its mean vector μ_i and covariance matrix $\sum_i$, where i denotes the i-th state;
- This multi-single Gaussian distribution is the to Σi defined by its mean vector μi and covariance matrix, where, i represents the i-th state;
- The state transition of the model obeys the property of first-order Markov, that is say, the state transition only depends on the choice of the previous state, having nothing to do with the given states previously;
- Any possible transfer from one chord to another is allowed, so the probability of state transition can gained by traverse way.

② Establish keynote-dependent HMM

In Western music, the tone is a very important attribute, but the tone and chord relationship is very close. Harmony on the basis of the tone, the chord sequence is a single chord linking according to certain rules. Therefore, the tone of known piece of music can provide a very valuable chord information. For example, if the tone of a known piece of music is in C major, we can infer that the C major chord, the F major chord, and the frequency of the G major chord. This is because the three chords in C major, respectively, corresponding to the tonic, under the tone is the sound. F # minor chord, and Ab major chord in C major harmony does not work, so it will not appear.

The keynote and tempo information of music can be extracted while using MIDI music files to obtain the script and features. Therefore, by the extracted keynote information, the keynote-dependent HMM can be established based on the above HMM.

- In Western music, it is very rare that the keynote of a music changes. So, we assume that one music have only one keynote, and the starting keynote as the keynote of the whole music;
- Define the major tone and the minor tone for each pitch. There are 12 kinds of pitches, so the numbers of the major tone and the minor tone are both 12. It needs to establish 24 different keynote models, that is to say λ_k, $1 \le k \le 24$;
- According to the different types of music, establish more keynote-dependent models.

The use of labeled training data from the MIDI file, we first trained the model to estimate the parameters of the model. Once we get the parameters of the model - the probability of initial state, the state transition probability matrix, mean vector and covariance of each state vector matrix - chord identification of an unknown input.

C. Keynote Estimation

Use Viterbi decoding to estimate the keynote and recognize the chord, the procedure is shown in Figure 3.

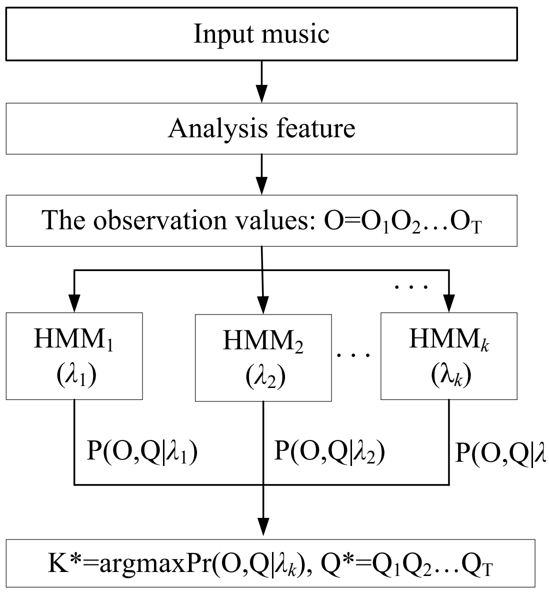

Figure 3. The procedure of keynote estimation

- Give the unknown music as input, and use the analyzer to analyze the 6-dimensional tonal centroid feature, then the sequence of the observations O={O_1 O_2 … O_T} can be gained;
- Do Viterbi decoding, and calculate the score of all 24 key-dependent models, record as P(O, Q|λ_k), $1 \le k \le 24$;
- Choose keynote-dependent model with the highest score, as shown as in (3).

$$K^* = \arg\max_{1 \le k \le 24} P\left(O, Q \middle| \lambda_k\right) \tag{6}$$

Thus the keynote of the music is estimated, while the optimal state path Q={Q_1 Q_2 … Q_T}is obtained, which is the chord sequence with frame-level precision.

III. TRAINING PROCEDURE OF CHORD RECOGNITION ALGORITHM

A. Labeled Training Data

Chord recognition system presented in this paper is first of all from the MIDI file has been marked chord information, and synthetic music audio from the same MIDI files to extract the feature vectors we need. Then the two parts of data for training the key-dependent the HMM. Secondly, in previous studies based on taking into account the importance and impact of the tone in the music, the tone of the music to the chord recognition system. The tune used in this article as equal temperament, define 24 tone, and therefore need to build 24 Key-dependent HMMs.

Hidden Markov Model is a model of a supervision order to train a supervised model, we need a lot of audio files, and these audio files corresponding to tag files contain the chord boundaries and chord names. To make this laborious process automatically, we use the feature data to tag files, while the establishment of the audio data. To achieve this aim, we will be a signature file to convert the file of another format, this format can be used as the input of the chord analysis tools. Then, the chord analyzer and sound analysis, root information and the tag name to output to a file. So that we can extract chord information from this file. Finally, when training HMMs, we use the chord sequence as a script.

The realization of the chord recognition system is divided into two parts. The first part is a training chord recognition system, the establishment of key-dependent HMMs. The second part is how to use chord recognition system based on the key-dependent HMMs for chord recognition and tone extraction.

Before testing the algorithm, the key-dependent HMM models need to be trained firstly.

As the dependence of the test model on the training data, in this study, the training data sets with two different types of music are chosen. One is made up of 239 classical music files, and the other one contains 120 music by Backstreet Boys download from Internet.

From synthesized MIDI music files, audio data, the data set length of time in classical music is 15.3435 hours, 252 588 characteristics. Beatles music data sets the length of time of 9.4756 hours, the characteristics of 155 889. Figure 5.1 shows the distribution of classical music dataset chord with the tone. Figure 5.2 indicates the distribution of the chord and tone of the Backstreet Boys music dataset.

The training procedure is as shown as in Figure 4.

- By harmony analysis, get the marked chords from MIDI music files, those are the script files;
- Synthesize the music audio in the same MIDI music file, and then extract the feature vectors from the synthesized audio;
- For each tone to establish a key-dependent HMM, a total establishment of 24 key-dependent HMMs;
- Set the feature vectors and script files as input to train the keynote-dependent HMM.

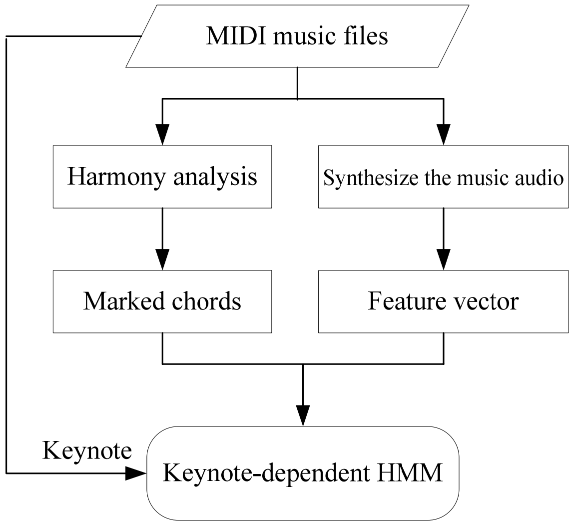

Figure 4. The training procedure

B. Parameter Estimation

1 Preprocessing of the training data

Chord recognition system is the key-dependent model, it is necessary for each key-dependent model should estimate its parameters. Therefore, in estimating the model parameters to preprocess the training data set. First classify the training data set, according to the tone is divided into 24 sub-training data sets, respectively, for the training data set of 24 key-denpendent model. Then, in order to be able to accurately estimate the parameters of the model, to avoid statistical error in the estimated parameters in the existing training data set, we carry out some operation to expand the amount of data of each sub-training data set to ensure that the sub-set of training data between the amount of data balance.

Specific preprocessing steps are as follows:

- Alignment handling of script files and signature files. That is, each frame features in the script file corresponding to the chords in the script file.
- On the outcome document of the script file and the alignment of the signature file in accordance with the tone classification, each of the key-dependent model training data set.
- Pairs training data set to be expanded. The data of all major or minor are shifted to a transfer of the data set. For example, in the training data set on the model of the C major expansion, we will all other major model training data set to pan. Left pan two in the training dataset D major chord names, such as the D major chord pan two into C major chord, and then the pan after the training data set into the C major model training data set. Expansion of the C major model training data set obtained. Expanded, the amount of data for all 12 major model training data sets are the same.

Minor model training data set to do the same treatment.

2 State transition probability estimates

After pretreatment model the training data set, we first estimate the state transition probability. Estimation model for the state transition probability matrix, we can observe that the transfer to each state to its probability is the largest. This is because the music a chord sustained time is greater than the frame rate when the feature extraction. In other words, when the extraction of the characteristics of two adjacent frames, the characteristics of the two frames are in the same period of chords.

In addition, the music of the C major, C major chord, the F major chord and the G major chord is the tonic, under the tone is the sound. F # minor chord, and Ab major chord does not appear. Tonic - subdominant in this kind of music - is a sound relationship will be reflected in the state transition probability. From the C major chord in C major state of transition probability, the state transferred to the F major chord and the G major chord probability than the probability of transfer to other chords.

3 Estimate the distribution of observations

The distribution of the values observed in a multi-single Gaussian distribution, defined by the values observed in the mean vector and covariance matrix. To preprocess the training data set, each key-dependent model of the training data set is balanced, but in each sub-training data set, the difference between the number of characteristics of the different chords, resulting in estimated a statistical error of the mean vector and covariance. Therefore, in each sub-training data set, we use similar pretreatment methods on the training data set, pairs training data set processing.

The approach is as follows:

- Pairs of training data set to analyze the characteristics of each chord corresponding to the classified.
- The characteristics of all the major chords in accordance with the relative position of the 12 law of averages pan, after the translation feature merged into a single chord set of features. For example, when estimating the mean vector and covariance of the C major chord, the characteristic cycle of C # major chord left one, the characteristics of the D major chord cycle left two, and so on, all the rest of the major chord the characteristics of cyclic translation of the corresponding number of bits. 12-chord feature set is the same size. Do the same treatment for the less chord 12 of the 12 small chords and classical music.
- 12-dimensional chroma feature vector is converted into the tone of the 6-dimensional feature vector from the heart.

Pairs of training data sets and data processing, we can calculate the mean vector and covariance matrix of each of the key-dependent model. As before we assume that whether it is 12-dimensional chroma feature vector or a 6-dimensional tonal centroid vector between the dimension and the dimension are mutually related, so the covariance matrix is strong diagonal, we only need to find the diagonal value can be.

IV. THE TEST RESULTS OF OUR ALGORITHM

To test the chord recognition algorithm, two types of music were used to test the performance of our chord recognition system, which are the popular music from Backstreet Boys and the Classical music. The test set of Classical music is composed of the “D major canon” from Pahaibeier and the "E flat major big beautiful Waltz." from Chopin. The popular test set is composed of two pieces of midi music from Backstreet Boys, “I want it that way”(C major) and “As long as you love me” (G major). These two test sets were excluded from the training set for our model. The details of the test set are shown in Table 1.

The tonal centroid vector is used by our recognition algorithm as the feature, and the recognition results with the key-dependent model and the key-independent model under two different training sets are shown in Table 2, respectively.

TABLE I.

THE MAJORS AND FRAMES OF TEST SETS.

Test set	Test music	Major	Frames
Classical music	canon	D major	441
	big beautiful Waltz	E flat major	1634
Backstreet music	I Want It That Way	C major	1152
	As Long As You Love Me	G major	1163

TABLE II
THE COMPARISON BETWEEN RECOGNITION RESULTS USING THE KEY-DEPENDENT MODEL AND THE KEY-INDEPENDENT MODEL

Training set	model	Test set（Precise%）			
		Classical music		*Backstreet Boys music*	
		canon	big beautiful Waltz	I Want It That Way	As Long As You Love Me
Classical music set	*Key-dependent*	60.08	65.59	57.45	62.14
	Key-independent	57.44	63.24	58.71	60.97
Backstreet Boys music set	*Key-dependent*	65.50	77.81	64.90	73.87
	Key-independent	64.62	75.13	63.47	72.29

From Table 2, it can be found that the result of key-dependent model is superior to the result of key-independent model. The reason is that the connectivity of chords is based on the major, which means that the known major of the music is helpful to the evaluation of the next possible chord in the music. Furthermore, the results in Backstreet Boys test set are better than the results in Classical music set, and the results of the two models in Classical music set are all better. The reason is there are 4 different styles of music in the Classical music training set, whereas there is only one style of music in Backstreet Boys training set. Therefore, the robustness of model under Classical music set is higher than the one under Backstreet Boys music set.

Table 2 shows, the six-dimensional tonal centroid vector as the feature recognition results are better than the 12-dimensional chroma vector. Tone from the heart vector is 12-dimensional chroma vector is mapped to a specific interval relations (five-tone ring, junior tone ring, three tone ring). In Western music, the interval relations defined most of the chords in the recognition of chords, the tone from the heart vector robustness is better than the chroma vector.

In Table 2, the test results obtained by the Beatles test set Bealtes model is better than the results on the model of classical music. Classical music test set in the classical music model and Bealtes of model test results are good. This is a classical music training set, a total of four composers, which is 4 style music, The Beatles training to focus on only one style of music. Therefore, the robustness of the models of classical music better than Bealtes model. However, we also found that the test set of classical music in the Beatles on the model test results are better than the results on the model of classical music. This is because the training set of classical music, the four composers each composer's music is too small (such as Haydn's music is only 4), such training model Bealtes model accurately.

V. CONCLUSIONS

Firstly, MIDI music files to the training data - the script and characteristics. MIDI music files, not only can automatically get the script, synthetic audio, feature extraction, but also be able to get tone information. Thus establish a basis for the introduction of the establishment of the HMM tone information. Secondly, in the establishment of the HMM, we take into account the importance of the tone in the music tone information into the HMM, the key-dependent model. Finally, the Viterbi decoder can get tone information and the optimal chord sequence.

A chord recognition algorithm based on key-dependent HMM is proposed in this paper. Firstly, it extracts the scripts and features from the training data composing of Midi music files, and refines the information of major, which provides the basis for the construction of HMM. Secondly, taking account of the importance of major in the music, the key-dependent model is built with the major information through HMM. Lastly, the major class and the optimal chord sequence can both be obtained via Viterbi decoding. The experiments had shown that the key-dependent model was capable to identify the major class and improve the precise of chord recognition.

Our chord recognition algorithm has built 24 models for each type of music. However, building the same set of models for each type of music will be a heavy task owing to the large number of music types. Therefore, building the key-dependent model without adapting the music type will be the next thing that will be considered in our future work.

REFERENCES

[1] D. Deutsch, "The psychology of music," 2nd ed. San Diego, CA: Academic, 1999.

[2] B. Su, S. Jeng, "Multitimber chord classification using wavelet transform and selforganized map neural networks," the IEEE International Conference on Acoustics, Speech, and Signal Processing(ICASSP), Salt Lake City, USA, 2001, pp.3377-3380.

[3] C. Giordano, P. Francois, and B. Jean-Pierre, "Recognizing chords with EDS: Part One," CMMR 2005, 2006, pp.185–195.

[4] T. Yoshioka, T. Kitahara, K. Komatani, T. Ogata and H. G. Okuno, "Automatic chord transcription with concurrent recognition of chord symbols and boundaries," Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR), Barcelona, 2004, pp100-105.

[5] A. Sheh and D. P. Ellis, "Chord segmentation and recognition using EM-trained Hidden Markov Models," In Proceedings of the International Symosium on Music Information Retrieval, Baltimore, MD, 2003.

[6] J. P. Bello and J. Pickens, "A robust mid-level representation for harmonic content in music signals," Proc. 6th Int. Conf. Mus. Inf. Ret., 2005, pp.304–311.
[7] E. Gomez and P. Herrera, "Automatic extraction of tonal metadata from polyphonic audio recordings," Audio Engineering Society, London, 2004.
[8] G. Y. Li, H. W. Zheng and C. P. Hu, "Analysis of Ttime-frequency domain characteristics of string music chord sound based on wavelet packet technique," Journal of Kunming University of Scinece and Technology (Science and Technology),Vol.28 No.4 Aug.2003, pp.36-39.
[9] J. Y. Sun, H. F. Li, and L. Lei, "ANN based music chord perception," Abstract Book of the 10~(th) National Conference on Man-Machine Speech Communication (NCMMSC'2009) and International Symposium on Speech and Language Processing, Urumqi, Xinjiang, China, Aug. 2009, pp.307-312.
[10] H. R. Yan, "Research on relational Markov networks and its application in social networks," Beijing: Beijing Jiaotong University, 2010.
[11] C. Harte, M. Sandler, "Automatic chord identification using a quantised chromagram," in Proc. AudioEng. Soc., Spain: 2005, paper 6412.