Looking for New, Not Known Music Only: Music Retrieval by Melody Style

Fang-Fei Kuo
Dept. of Computer Science and Information Engineering
National Chiao Tung University
HsinChu, Taiwan, ROC
ffkuo@csie.nctu.edu.tw

Man-Kwan Shan
Dept. of Computer Science
National Cheng Chi University
Taipei, Taiwan, ROC
mkshan@cs.nccu.edu.tw

ABSTRACT
With the growth of digital music, content-based music retrieval (CBMR) has attracted increasingly attention. For most CBMR systems, the task is to return music objects similar to query in syntactic properties such as pitch and interval contour sequence. These approaches provide users the capability to look for music that has been heard. However, sometimes, listeners are looking, not for music they have been known, but for music that is new to them. Moreover, people sometimes want to retrieve music that "feels like" another music object or a music style. To the best of our knowledge, no published work investigates the content-based music style retrieval. This paper describes an approach for CBMR by melody style. We proposed four types of query specification for melody style query. The output of the melody style query is a music list ranked by the degree of relevance, in terms of music style, to the query. We developed the melody style mining algorithm to obtain the melody style classification rules. The style ranking is determined by the style classification rules. The experiment showed the proposed approach provides a satisfactory way for query by melody style.

Categories and Subject Descriptors
H.2.8 [**Database Management**]: Database Applications – *Data mining;* H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *Query formulation, retrieval models, search process;* H.3.7 [**Information Storage and Retrieval**]: Digital Libraries – *Systems issues;* H.5.5 [**Information Interfaces and Presentation**]: Sound and Music Computing – *Methodologies and techniques;* J.5 [**Computer Applications**]: Arts and Humanities – *Performing arts.*

General Terms
Algorithms, Design, Experimentation, Human Factors.

Keywords
Content-Based Music Retrieval, Music Style Mining, Query by Melody Style, Music Classification.

JCDL'04, June 7–11, 2004, Tucson, Arizona, USA.

1. INTRODUCTION
Music information retrieval (MIR) has become an increasingly important field of research in recent years. In traditional MIR systems, the query is based on text-based metadata. The content-based music retrieval (CBMR) allows user to query by music content instead of metadata.

Much work has been done on the development of CBMR. Query by humming or singing[7][8][11][13][14][18] are common approaches for retrieval from acoustic input. The queries were melodies hummed or sung by the user, and were transcribed into symbolic MIDI format. Query by tapping is another query method that takes the beat information for retrieval[10]. Recently, several researchers have explored polyphonic content-based music retrieval7[15][16]. The polyphonic music retrieval techniques are more suitable than monophonic music retrieval for retrieving performance data and query by polyphonic input.

Main goals of the previous CBMR researches are to return the music objects that are similar to the query in pitch, interval contour or rhythm. Moreover, these CBMR approaches provide users the capability to look for music that they have been heard. However, sometimes, listeners are looking, not for music they have been known, but for music that is new to them. Moreover, people sometimes want to retrieve music that "feels like" another music object or a style.

To look for new music that we haven't listened, the approaches of query by humming, singing, or tapping is helpless. It is necessary to develop the technique for query music by melody style.

Music style implies the human perception of music, which is the feature that people often utilize to classify music. Though text-based metadata, which records the text description of music style, can be utilized for melody style query, it should be annotated manually. Furthermore, sometimes user may wish to query mixed style. For example, the users may want to retrieve music mainly sounds like Chopin and a little Bach. The returned music objects should be more similar to Chopin style but also have a little feeling of Bach style.

The purpose of our research was to investigate the technique for content-based music retrieval by melody style. There are several issues about our work:

(1) To develop the methods for the specification of query style.

(2) To determine the appropriate feature for music style and its representation.

(3) To discover the description of melody style.

(4) To measure the degree of relevance between the music object and the query style.

For the first issue, we present four types of query specification for query style. For the second issue, the basic elements of music consist of melody, harmony, rhythm, and so on. Above all, melody is the most memorable aspect of music. Accordingly, we concentrated on the melody style and utilized chord as the melody feature for retrieval by music style. For the third issue, we develop an algorithm to discover the common characteristics from the music of the same style and find the discriminating patterns between the music of various styles. The melody styles are described by the discovered set of style rules. For the last issue, the discovered set of style rules is used to rank the music objects.

Our work is useful in many aspects of applications. For example, to help physiotherapist for seeking music that will motivate a patient, to help film director for seeking music conveying a certain mood[9], to help restaurateur for seeking music that targets a certain clientele. Query by melody style provides users the capability to find music with style similar to what users like.

This paper is organized as follows. Section 2 give a brief review of previous work related to content-based music retrieval and music style discovery. In section 3, we present the music style retrieval model. Section 4 describes our proposed methodology. The experiment and result of performance analysis is described in Section 5. Section 6 concludes the paper.

2. RELATED WORK

Much research has been done on the development of the content-based music retrieval technology. Query by humming or singing is a common approach for query by acoustic input[7][8][11][13][14][18]. Ghias et al.[7] introduced a query by humming system. The query input was converted into a melodic contour and the contour was matched against the music in the library by approximate string matching. McNab et al.[14] presented a CBMR system that accepted singing or humming queries. They investigated people's singing accuracy and suggested that the music transcription should adapt user's tuning. In Tseng's research[18], key melody extraction is used for query suggestion and effective retrieval, where the key melodies are representative fragments of music. To allow queries in any key levels and match approximately, the pitch profile encoding and n-note indexing techniques were used respectively. Kline et al.[11] developed approximate matching algorithms make better use of both pitch and duration information, which improved results when the users have relatively little music experience or ability. Lu et al. [13] proposed a new melody representation and hierarchical matching method for query by humming system. The melody representation is a combination of pitch contour, pitch interval and the duration. Jang et al.[10] presented a new query paradigm, which allows user query by tapping. Melodies are transformed into the time vectors that contain the beat information. Hu et al. [8] compared the performance of several retrieval algorithms. The types of query include humming, singing and whistling. In 7, Chen et al. investigated the music content representation and retrieval techniques. They proposed music segment as a music content representation, which consists of both melody and rhythm information.

Several researchers have explored polyphonic content-based music retrieval7[15][16]. Doraisamy et al. 7 proposed the polyphonic music indexing using pitch and rhythm information. In [16], a probabilistic model is proposed for retrieving performances that include large number of variations in performing a melody and accompaniment. Pickens et al. [15] proposed harmonic description which contains the information from all chords, and combined with Markov method to model music document and query.

Though the aim of this work is melody style retrieval rather than melody style classification, several works on music genre classification that are related to our work are described as follows. The work developed in MIT Media Lab.[1] employed hidden Markov model to model and classify the melodies, which were represented as a sequence of absolute pitches, absolute pitches with duration, intervals and contours. Another research in CMU used the naïve classifier, linear and neural network respectively to recognize music style for interactive performance systems [5]. Thirteen statistical features derived from MIDI are identified for learning of music style. In [19], the music genre classification algorithms aimed at audio signals were explored. They proposed features for representing the musical surface and rhythmic structure and classified by statistical pattern recognition classifier.

3. MUSIC STYLE RETRIEVAL MODEL

Before the description of the proposed approaches of music style retrieval, we first formalize how the music style is modeled.

Definition 1 A music object O is represented as $O = O(M, F, R)$ where

- M is the raw music data, for example, an MIDI file.
- $F = \{f_i\}$ is a set of low level music features associated with the music object.
- $R = \{r_{ij}\}$ is a set of representations for a given feature f_i.

Style usually refers to collections of data. Style is a concept description that generates descriptions for *characterization* and *discrimination*. Characterization refers to the common patterns of a given collection while discrimination denotes the comparison among collections. Therefore, the music style involves both the characterization of music patterns for each collection of music object and the discrimination of music features among collections of music objects.

Definition 2 The *music style* T is modeled as $T = D(C(G(O)))$ where G is the taxonomy of the music objects, C is the characterization function, D is the discrimination function.

For example, the taxonomy of music objects may be classified according to the composer. For the folk song, the taxonomy may be classified according to the peoples. For the Western music, the taxonomy of music objects may be classified according to the eras of history of Western music, namely, the Baroque, the Classical, the Romantic and the Modern era.

For the taxonomy of Western music, the music shares aspects of style with other pieces written at roughly the same time. In the Baroque era, melodies are ornate and often make use of dramatic leaps. Repetition and simple binary and ternary forms provide the basis for musical structure. Rhythms are often derived from dance rhythms. Harmony is based on major/minor tonality, and dissonances become more common. The music style of Classical era is reflected in simple texture (homophonic textures became the standard while contrapuntal texture was used sparingly), simple melodies (melodies usually fall into even phrases, and often were organized into symmetrical "question and answer" structures) and simple, rational forms (simple two- and three-part forms became the essential building blocks of all Classical forms, especially the Sonata Allegro form). In the Romantic era, the melodies are longer, more dramatic and emotional. Moreover, Tempos are more extreme. Harmonies are fuller, more dissonant. In the Modern era, melodies can be long and abstract or reduced to small gestures. Form can be controlled to an almost infinite degree, or it may be the result of improvisation and chance.

Definition 3 The music style retrieval is modeled as $S = S(T, O)$ where S is the *ranking function* which measures the similarity between a given music object O and a specific music style T.

4. METHODOLOGY

4.1 Query Specification

The style query can be described in many ways. In our work, we proposed four types of query specification for music style query as follows.

(1) Query-by-music-group (QBMG): The user specifies the query style by selecting a group of music from the example music. The set of example music are randomly generated by the system. Therefore, the common style of the selected music group is what the user wish to retrieve. The constitution of these query examples can be regarded as a new, user-defined music style.

(2) Query-by-music-example (QBME): This is similar to query-by-music-group with the exception that only one example is selected. In this way, the user can retrieve the music with style similar to the query example.

(3) Query-by-taxonomic-style (QBTS): An example is to retrieve the music with Baroque style.

(4) Query-by-taxonomic-style-combinations (QBTSC): For instance, to retrieve the music with both Baroque and Romantic styles. In this way, the combination of these styles can be viewed as a new style.

To process these four types of query, Figure 1 shows the flowchart of our approach. The kernel is the feature extraction and feature representation module. For each MIDI file in the music digital library, after the offline processing of the feature extraction and representation, the corresponding representations are stored in the library. The feature extraction and representation modules firstly process each of the four types of query issued by the user. For the query of type QBME, the representation of the extracted feature is then evaluated against each of the corresponding representation of MIDI files in the library and the ranking list is generated. For the query of type QBMG, QBTSC, or QBTS, the style patterns generated from the query are evaluated against each of the corresponding representation of MIDI files in the library and the ranking list is generated. The style patterns are generated by characterization and discrimination from the music set specified in the query. For QBMG, the music set is the selected group of music. For QBTS and QBTSC, the music set is the music corresponding to the specific taxonomy of music.

4.2 Feature Extraction

Music is usually polyphonic, in which two or more notes sound simultaneously. Since we focus on the melody style, it is necessary to extract melodies from MIDI files. We have proposed the melody extraction method for this task [12]. This method considers the information of instrument, volume and highest pitch of MIDI. Then, the proposed chord assignment algorithm extracts chords from the melody [12]. The chord assignment algorithm is a heuristic method based on harmony and music theory. Sixty common chords are chosen as the candidates. For each melody, the algorithm first decides length of the sampling unit used for music segmentation. The chord candidates are scored for each sampling unit, and the highest one is assigned to the sampling unit. The algorithm may assign a set of chords (chord-set) to a sampling unit while chord with the highest score is not unique. Output of the chord assignment algorithm is a sequence of chord-sets, and the chords are represented in Roman numerals such as Ⅰ, Ⅲmaj, Ⅵm7 for key invariant. For more detail explanation of the chord assignment algorithm, please refer to [12].

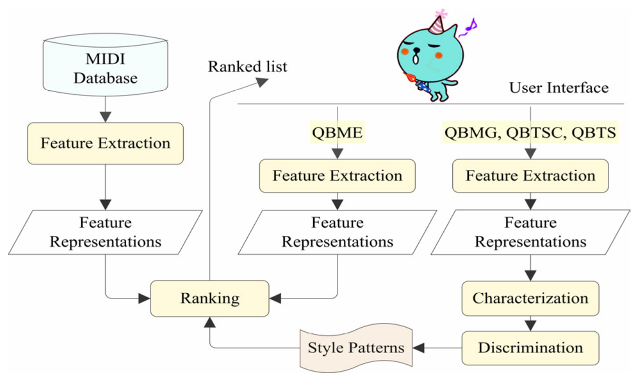

Figure 1. Flowchart of proposed approach.

4.3 Feature Representation

After feature extraction, there are three different representations for the chord feature as follows:

(1) Set of chord-sets: music is represented as a set of items, where each item is a chord-set.

(2) Set of bigrams: music is represented as a set of bi-grams of chord-sets. A bi-gram is an adjacent pair of chord-sets extracted from a sequence of chord-sets. Therefore, a melody with n units consists of $(n\text{-}1)$ bi-grams.

(3) Sequence of chord-sets: music is represented as a sequence of chord-sets. In this way, a melody with n units is actually an n-gram.

4.4 Query Processing

4.4.1 Query-By-Music-Group

As stated in Section 3, the music style involves both the characterization and discrimination of music features. Therefore, to process this type of query, there are three major steps.

(1) The first step is to discover the common characteristics of the selected group and the unselected group of music examples respectively.
(2) The second step finds the discrimination between the characteristics of these two groups. The result of this step is a two-way classifier.
(3) At last, a ranking function is employed to measure of degree of relevance between a music object and the query style based on the two-way classifier. Given the ranking function, all the music objects in the library are evaluated and a ranking list is produced and output to the user.

Characterization

The first step takes the features of the selected group and the unselected group as input respectively. Frequent pattern mining technique is employed to derive the common properties and the interesting hidden relationships between chords and melody styles from music of the same group. Two frequent pattern mining methods are utilized with respect to the representations of the melody style feature.

If the melody feature is represented as the set of chord-set or the set of bi-grams, the concept of frequent itemset in the association rule mining is utilized [2]. In the terminology of association rule mining, *support* of an item-set is defined as the percentage of transactions which contain this item-set. Given the *minimum support* specified by the user, an item-set is *frequent* if its support is larger than the minimum support.

In our approach, the transaction database for the selected group (or the unselected group) consists of the features of music belonging to the selected group (or the unselected group). Each transaction corresponds to the set of chord-sets of a specific music. In other words, a chord-set is corresponding to an item in the terminology of association rule mining. The frequent item-set denotes the set of chord-sets which are accompanied together with the melodies of most music in the selected group. For example, assume that there is the frequent item-set {{I}, {V, ⅥIm7}, {V}} for the lyric-style music, this represents that the melodies of a great part of lyric-style music consist of chord-set {I}, {V, VIm7} and {V} together. The same concept is applied for representation of set of bigrams. That is, a bigram of chord-sets corresponds to an item.

If the feature of melody style is represented as the sequence of chord-sets, to find the common characteristics of music of the same group, we propose a new type of pattern – *frequent consecutive sequential pattern*. The concept of frequent consecutive sequential pattern is modified from that of sequential pattern [2] in sequence data mining techniques. The consecutive sequential pattern is continuous, which differs from the original sequential pattern. A consecutive sequential pattern is said to be contained in a transaction if the pattern is a consecutive subsequence of this transaction. For example, the consecutive sequential pattern ({V, ⅥIm7}, {V}, {I, III, Vim7}) is contained in the transaction ({I}, {V, ⅥIm7}, {V}, {I, III, Vim7}) while ({V, Ⅵ m7}, {I, III, Vim7}) is not. The support of a consecutive sequential pattern is defined as the percentage of transactions which contain it. Given the minimum support specified by the user, a consecutive sequential pattern is frequent if its support is larger than the minimum support. We modified the join step of the Apriori-based sequential mining algorithm to find frequent consecutive sequential pattern.

Discrimination

The frequent patterns indicate the common properties of the music objects belong to the same style. However, it is not enough to discriminate one style from others only by the frequent patterns. In generally, people recognize a music style not only by the characteristics of itself, but also by the differences between this style and others. Discrimination tries to find the discrimination among characteristics of music group. The result of the discrimination for taxonomy of music groups is a *melody style pattern set* which consists of *melody style rules*.

Definition 4 The *melody style rule* r is of the form $l \Rightarrow y$, where y is a music group corresponding to a melody style and l is the characteristics of y which may be a frequent set of chord-sets, a frequent set of bigrams or a frequent consecutive sequential pattern.

Definition 5 The *melody style pattern set* is an ordered set of melody style rules. Format of the melody style pattern set is $< r_1, r_2, ..., r_n, default\_class >$, where each melody style rule r_i is ranked by the confidence. Given the set of music and the taxonomy, the *confidence* of a rule r_i is the percentage of music objects satisfying the characteristics of r_i belong to the music group y_i.

In our work, if the type of characteristics is the consecutive sequential pattern, then the feature f of a music object satisfying the characteristics l if f is contained in l. If the characteristics is the set of chord-sets or the set of bigrams, then the feature f of a music object satisfying the characteristics l if f is a subset of l.

The melody style pattern set may be regarded as a classifier which is learned from the given taxonomy of music objects and corresponding characteristics. It can be used to classify music of unknown group. To classify the music object, the first rule that satisfies the music is used to classify it. If there are no rules satisfying, the music is classified according to the default_class. Figure 2 shows an example of melody style pattern set.

We proposed a melody style classification algorithm in [12], which is based on frequent patterns to differentiate the melody styles. In this work, we employ the classification algorithm to generate classification rules and regard the rules as the melody style pattern set. The characteristics of our proposed melody style rules consist of frequent set of chord-sets, frequent set of bigrams and frequent consecutive sequential pattern. Moreover, some music styles may contain more rules of lower support while some styles have fewer rules of higher support. In other words, the appropriate values of minimum support differ from each other. The rules built by our classification algorithm consist of multiple types of characteristics and the minimum support of each rule may differ. Our algorithm uses five-fold cross-validation to determine the appropriate minimum supports. For more detail of the music style classification algorithm, refer to [12].

Ranking Function

After the generation of the melody style pattern set respective to the style of query music group, the similarity between the music in digital library and the query style is evaluated as the way of classifying the music data. As stated in the previous subsection, the melody style rule in the melody style pattern set is ordered according to the confidence. The confidence implies the degree of membership where the characteristic of the rule belongs to the style. Hence ranking of the music data is decided by the confidence of the first rule that satisfies the music data.

If the first matched rule for music in library does not belong to the style of the selected group, the music in the library is not a qualified answer. Otherwise, the confidence this rule is regarded as the ranking measure for this music in library. Take the example of Figure 2, if the sequence of chord-sets of a music object in library is { Ⅱ7 Ⅴ Ⅲ Ⅱ Ⅴ Ⅰ Ⅶ}, it matches the third pattern style rule. The ranking score of this music object respect to the query group is 0.6.

Set: { Ⅰ, Ⅲ, Ⅳ7} → style 1, conf = 0.9
Bigram: {(Ⅴ Ⅰ), (Ⅴ7 Ⅶ)} → style 2, conf = 0.75
Sequence: (Ⅴ Ⅲ Ⅱ Ⅴ Ⅰ) → style 1, conf = 0.6
Bigram: {(Ⅰ Ⅱ), (Ⅳ7 Ⅴ), (Ⅱ Ⅵ)} → style 1, conf = 0.57
Default_class: style 2, conf=0.55

Figure 2. An example of melody style pattern set.

4.4.2 Query-By-Music-Example (QBME)

Query-by-music-example allows users to query similar style music by an example of music rather than by a group of music. For QBME, we do the style matching for the music in the library and query music directly. The style matching is measured based on the similarity of melody feature between the library and the query music. The result of QBME is a list of music ranked by the similarity.

As stated in section 4.2, the extracted chords are used as the feature of melody. Consequently, the melody style matching process becomes the similarity measurement of chord features. We first give the definitions for the feature representation of chord-sets.

Definition 6 Given two chord-sets u and v, the similarity $s(u, v)$, between them is defined as $s(u,v) = |u \cap v| \big/ \sqrt{|u| \times |v|}$, where $|u|$ is the cardinality of the set u, $\cap$ is the set intersection operation.

Definition 7 Given two sets of chord-set $U=\{u_1, u_2,\ldots, u_M\}$ and $V = (v_1, v_2,\ldots, v_N)$, the similarity constraint δ, and the similarity $s(u_i, v_j)$, $\forall\, i$, $1 \le i \le M$, $\forall\, j$, $1 \le j \le N$, a mapping between them is a one-to-one relation R_{set} from $\{1, 2, \ldots, M\}$ to $\{1,2, \ldots, N\}$, such that for each order pair (i, j) in R_{set}, $s(u_i, v_j) \ge \delta$.

Definition 8 Given two sets of chord-set $U=\{u_1, u_2,\ldots, u_M\}$ and $V = (v_1, v_2,\ldots, v_N)$, the similarity constraint δ, the similarity between U and V for a mapping R_{set}, $S'_{R_{set}}(U, V, \delta)$, is defined as

$$S'_{R_{set}}(U,V,\delta) = \sum_{\forall (i,j) \in R_{set}} s(u_i, v_j) \Big/ \sqrt{M \times N}$$

Definition 9 Given two sets of chord-set U and V and the similarity constraint δ, the similarity between U and V $S_{set}(U, V, \delta)$ is defined as

$$S_{set}(U,V,\delta) = \max_{\forall R_{set}} \{S'_{R_{set}}(U,V,\delta)\}$$

Example 1 Consider the following two sets of chord-set:

$$U = \{u_1, u_2, u_3, u_4\} = \{\{\mathrm{I}, \mathrm{V}\}, \{\mathrm{IV_m}\}, \{\mathrm{I}, \mathrm{IV}\}, \{\mathrm{II}, \mathrm{IV}, \mathrm{VI_m}\}\} \text{ and}$$

$$V = \{v_1, v_2, v_3\} = \{\{\mathrm{I}, \mathrm{IV}\}, \{\mathrm{II}, \mathrm{V}, \mathrm{IV_m}\}, \{\mathrm{V_{maj}}, \mathrm{IV_m}, \mathrm{II}\}\}.$$

Given the similarity constraint $\delta = 0.4$, the pairs of chord-set whose similarities are larger than or equal to δ consist of (u_1, v_1), (u_1, v_2), (u_2, v_2), (u_2, v_3), (u_3, v_1) and (u_4, v_1), and their similarities are $1/2$, $1/\sqrt{6}$, $1/\sqrt{3}$, $1/\sqrt{3}$, 1 and $1/\sqrt{6}$ respectively. $S_{set}(U,V,\delta) = 0.986$.

To find the similarity defined in Definition 9, we employed the Kuhn-Munkres algorithm (also known as Hungarian method). Given a weighted complete bipartite graph $G=(U \cup V, U \times V)$, the Kuhn-Munkres algorithm finds a matching from U to V with maximum weight. Such a matching from U to V is called an optimal matching.

For the representation of bigram set, the definition of similarity is similar to those of the set representation. The only exception lies in the similarity measure between two bigrams.

Definition 10 Given two bigrams x and y, where $x==u_1 \bullet u_2$, $y= v_1 \bullet v_2$, the similarity $s(x, y)$ between them is defined as

$$s(x,y) = \frac{|u_1 \cap v_1|}{\sqrt{|u_1| \times |v_1|}} \times \frac{|u_2 \cap v_2|}{\sqrt{|u_2| \times |v_2|}},$$

where $|u|$ is the cardinality of the set u, $\cap$ is the set intersection operation.

Example 2 Consider the following two bigrams:

$$x = u_1 \bullet u_2 = \{\mathrm{I}, \mathrm{V}\} \bullet \{\mathrm{IV_m}\} \text{ and}$$

$$y = v_1 \bullet v_2 = \{\mathrm{I}, \mathrm{IV}\} \bullet \{\mathrm{II}, \mathrm{V}, \mathrm{IV_m}\}.$$

The similarity $s(x,y) = 1/2 \times 1/\sqrt{3}$

Definition 11 Given two chord-set sequences $A = (a_1, a_2,\ldots, a_M)$ and $B = (b_1, b_2,\ldots, b_N)$, the similarity constraint δ, and the similarity $s(a_i, b_j)$, $\forall\, i$, $1 \le i \le M$, $\forall\, j$, $1 \le j \le N$, a mapping between them is a one-to-one relation R_{seq} from $\{1, 2, \ldots, M\}$ to $\{1,2, \ldots, N\}$, such that

(1) For each order pair (i, j) in R_{seq}, $s(a_i, b_j) \ge \delta$,

(2) For any two ordered pairs (i, j), (k, l) in R_{seq}, $[(j - l)=1]$ if and only if $[(i - k)=1]$.

Definition 12 Given two chord set sequences $A = (a_1, a_2,\ldots, a_M)$ and $B = (b_1, b_2,\ldots, b_N)$, the similarity constraint δ, the similarity between A and B for a given mapping R_{seq}, $S'_{R_{seq}}(A, B, \delta)$, is defined as

$$S'_{R_{seq}}(A,B,\delta) = \sum_{\forall (i,j) \in R_{seq}} s(a_i, b_j) \Big/ \sqrt{M \times N}.$$

Definition 13 Given two chord set sequence A and B, the similarity constraint δ, the similarity between A and B $S_{seq}(A, B, \delta)$ is defined as

$$S_{seq}(A,B,\delta)=\max_{\forall R_{seq}}\{S'_{R_{seq}}(A,B,\delta)\}.$$

Example 3 Consider the following two sequences of chord-set:

$$A=(a_1,a_2,a_3,a_4)=(\{\mathrm{I,V}\},\{\mathrm{IV_m}\},\{\mathrm{I,IV}\},\{\mathrm{II,IV,VI_m}\})$$ and

$$B=(b_1,b_2,b_3)=(\{\mathrm{I,IV}\},\{\mathrm{II,V,IV_m}\},\{\mathrm{V_{maj},IV_m,II}\}).$$

Given the similarity constraint δ = 0.4, the similarity between A and B $S_{seq}(A,B,\delta)=\left(1/2+1/\sqrt{3}\right)/\sqrt{4\times 3}$.

To compute this similarity measure, the algorithm is based on the dynamic programming strategy.

4.4.3 Query-By-Taxonomic-Style (QBTS) and Query-By-Taxonomic-Style-Combination (QBTSC)

QBTS allows users query music by system predefined taxonomic style. To process this query, preprocessing for the generation of melody style pattern set corresponding to the predefined taxonomic style is required. The music objects in the library are grouped according to this predefined taxonomy. If the taxonomy consists of *m* styles of music, then there are *m* groups of music in the library. The generation of melody style pattern set for these *m* groups of music is similar to that for QBMG. The only exception lies in the number of music groups. In QBMG, there are only two music groups, one for the selected group and the other for the unselected group. After the generation of music style pattern set for QBTS, ranking is of the same as that in QBMG.

For the query of QBTSC, the generation of music style pattern set is of the same as that for QBTS. Ranking is done by multiplication of the ranking scores respective to the styles specified in QBTSC.

5. EXPERIMENTS

We have evaluated the effectiveness of the proposed melody style mining approach. For more detail, please refer to [12] and [17]. In this paper, we focus on the evaluation of the performance of the proposed style query specification and ranking measures. We have implemented a music style retrieval system (http:// 140.113.215. 246) to perform the experiments. The music digital library contains four music styles of classical music – Baroque, Classic, Romantic and Modern style, each style contains fifty MIDI files. All MIDI files were gathered from the Internet. The Baroque style includes music of J. S. Bach, Vivaldi and Handel. The Classic style contains music composed by Haydn and Beethoven. The Romantic style includes music of Chopin and Brahms. The Modern style consists of music of Debussy, Ravel, Prokofiev and Saint-Saens. The music of Bach was downloaded from http://www.bachcentral.com. Beethoven and Brahms's music were downloaded from http://www.midi iofm.net. Chopin's music was acquired from the web site http://egalvao.com/chopin. The others were accessed from http://www.music-scores.com. For each file in the library, the melody extraction and chords assignment were performed. Figure 3 shows the snapshot of the query by music group while Figure 4 shows the results returned by the system.

We invited ten users whose backgrounds cover various levels of music training to perform the experiments. One user had learned guitar for several years, three had learned piano for a few years, one is the co-leader of the chorus, one is highly interested in classic music and the others don't have more music discipline besides the basic music courses in the school.

For each type of proposed query specifications, the users made three rounds of tests respectively. In each round of test, they made the query and gave scores to the music files in the result lists based on their perception of the style similarity between query and results. The users were requested to listen to all music files in the result list to ensure the reliability of the scores. There are seven levels of the score: -5, -3, -1, 0, 1, 3, 5, where the score 5 indicates the highly relevant and -5 indicates the highly non-relevant.

For the QBTS and QBTSC methods, users should know the characteristics of the Baroque, Classic, Romantic and Modern styles. To give users roughly knowledge about these styles, the system provided a brief introduction and some famous works for each style. Table 1 shows these representative works.

Table 1. Representative works for each style.

Style	Music title	Composer
Baroque	Cantata No.147: Jesu, Joy of Man's Desiring	J.S. Bach
	Invention in a minor, BWV 784	J.S. Bach
	Inven------tion in C major, BWV 772	J.S. Bach
	Messiah No. 7 Chorus: And he shall purify	Handel
	The Four Seasons: "Autumn" (Allegro)	Vivaldi
Classic	Trumpet Concerto in Eb, 3rd movement	Haydn
	Bagatelle No. 3, Op. 33	Beethoven
	Ruins of Athens Overture, Op. 113	Beethoven
	Moonlight Sonata Op. 27 No. 2, 1st	Beethoven
	Fur Elise	Beethoven
Romantic	Mazurka in Bm, Op. 33 No. 4	Chopin
	Mazurka in F#m, Op. 59 No. 3	Chopin
	Mazurka in Bb, Op. 7 No. 1	Chopin
	Etude in E, Op. 10 No. 3	Chopin
	Hungarian Dance No. 5	Brahms
Modern	Golliwogg's Cake-walk	Debussy
	Doctor Gradusad Parnassum	Debussy
	Serenade for the doll	Debussy
	Bolero	Ravel
	Carnival of the Animals: Elephant	Saint-Saens

The system generated random music lists for users to select the query example(s) for QBMG and QBME. There are twenty and ten music files in the query list of QBMG and QBME respectively. For QBTS, QBTSC and QBMG, the number of music in the result lists is twenty, and system returned ten query results for the proposed three similarity measures of QBME. As we have stated in the first section, music retrieval by style try to find the music which is similar to the query style. People wish to find something new, not something known. Therefore, it is not adequate to measure the performance by recall. We measure the performance only by precision and average scores given by the users. Precision is defined as

$$precision = N_{retrieved\_relevent} / N_{retrieved},$$

where $N_{retrieved\_relevant}$ is the number of relevant music retrieved and $N_{retrieved}$ is the number of retrieved music. The music is relevant if its score is larger than or equals zero. The average score is defined as

$$average\_score = \sum_{i=1}^{N_{retrieved}} Score_i \Big/ N_{retrieved} \ ,$$

where the $Score_i$ is the score of music i feedback by the user.

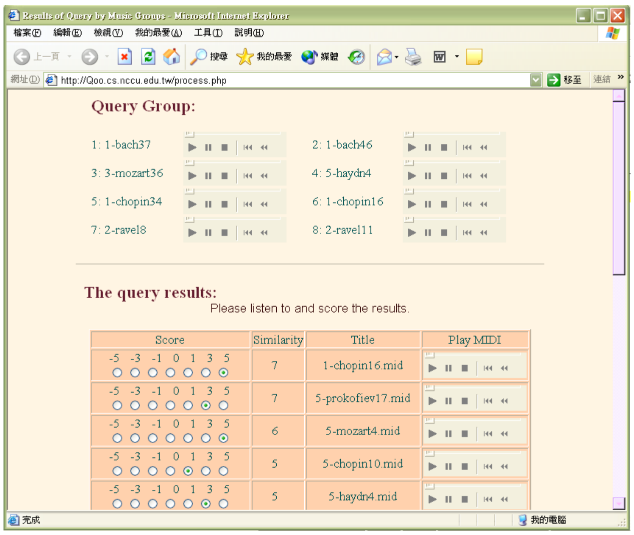

Figure 3. Snapshot of query-by-music-group.

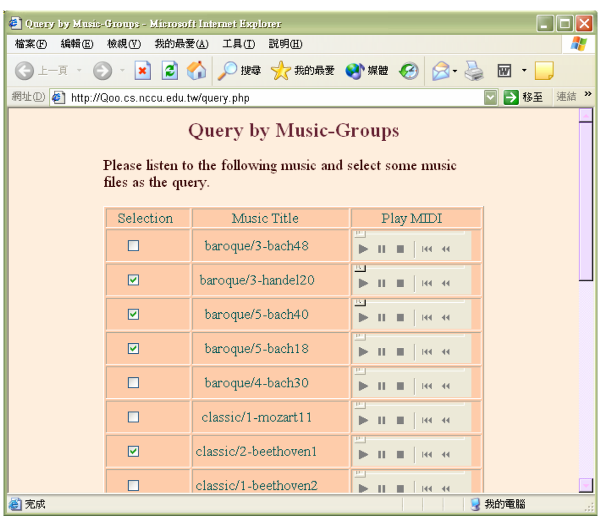

Figure 4. Snapshot of query result.

We calculate the precision and average score for each round of query of the users, and average the precisions and average scores of each user. The overall performance of each type of proposed query specifications and similarity measures is the average of all user's average precisions and average scores. Figure 5 shows the average precision and average score curves for the three similarity measures of QBME respectively. Both the average precision and average score curves are downward gradually. The average precisions range between 0.63 and 1, and the average scores range between 0.62 and 4.73. There are no significant differences among the set, bigram and sequence similarity measures, but in most case the bigram similarity performs better. In the following experimental results, we use the results of bigram similarity measure for QBME.

The precision and average score curve of the four types of query specification are shown in Figure 6. The range of precision of QBTSC and QBTS is between 0.86 and 0.91, QBMG is between 0.71 and 0.83, QBME is between 0.66 and 1. The range of average score of QBTSC and QBTS is between 2.26 and 3.27, QBMG is between 0.82 and 2.24, QBME is between 0.62 and 4.64. The precision curves of QBTSC and QBTSC are flat; QBMG and QBME are downward gradually. The average scores of all query specification types are tending downwards. The results show the QBTSC and QBTS perform better than QBMG and QBME, and the QBTS has higher average scores than QBTSC.

For the QBTSC and QBTS, the query is one or a combination of taxonomic styles, and the query of QBME and QBMG is one or a number of music files. This means that the scope of query style of QBTSC and QBTS is larger than that of QBME and QBMG. The slopes of the precision curves reflect this difference. There are more music files corresponding to the query style of QBME and QBMG, so the precision keeps high. On the contrary, the query style of QBMG is more specific and the slope of precision curve is larger; there is only one music file in the query of QBME, so its slope is largest. Furthermore, the users may be stricter while the query is more specific.

6. CONCLUSIONS

In this paper, we have proposed an approach for melody style retrieval. We proposed four types of query specification for melody style query. Query processing of these four types of query was presented. Query processing involves the steps of the feature extraction, feature representation, melody style pattern generation and ranking. The melody style pattern generation is an integration of characterization and discrimination. The performance measured by the precision indicates that the test users are satisfied by the result returned by the system.

Our work provides a new and effective way for retrieval in terms of music style rather than the syntactic features of music. Future research could provide other query methods such as query by selecting multiple styles, query by style example music and define the corresponding similarity measures.

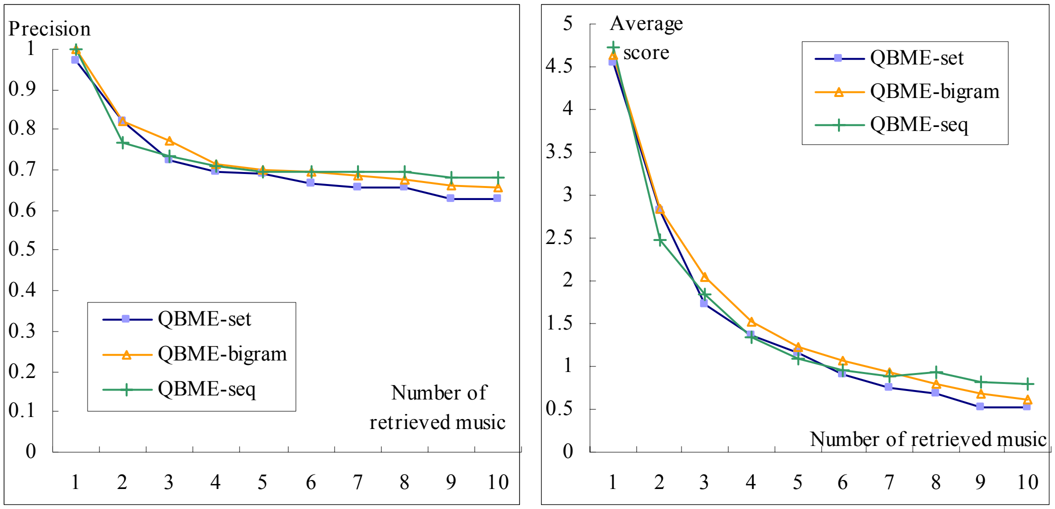

Figure 5. Average precision and score curves of QBME.

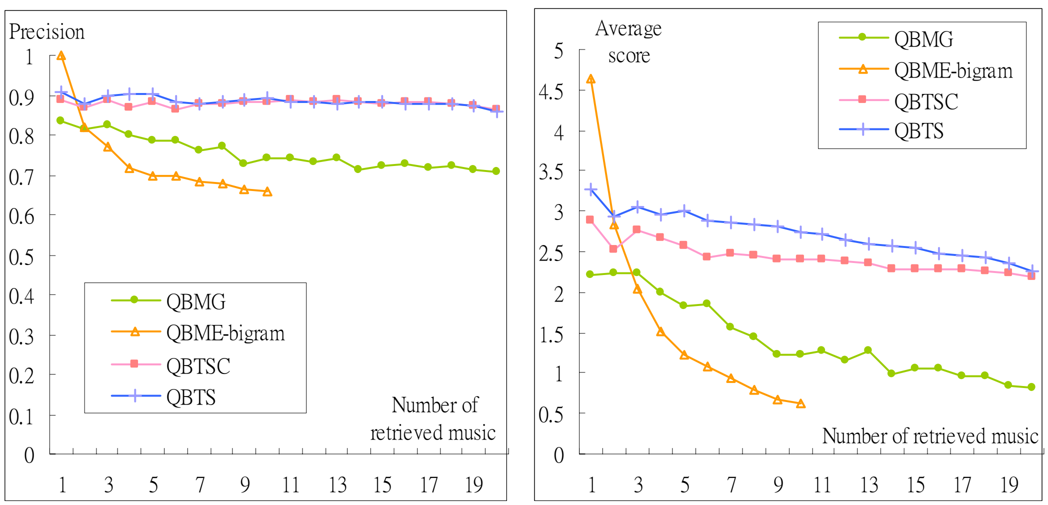

Figure 6. Average precision and score curves of all users.

7. REFERENCES

[1] Agrawal, R. and Srikant, R. Fast Algorithms for Mining Association Rules. In *Proceedings of the International Conference on Very Large Data Bases*, 1994.

[2] Agrawal, R. and Srikant, R. Mining Sequential Patterns. In *Proceedings of the International Conference on Data Engineering,* 1995.

[3] Chai, W. and Vercoe, B. Folk Music Classification Using Hidden Markov Models. In *Proceedings of the International Conference on Artificial Intelligence*, 2001.

[4] Chen, A. L. P., Chang, M., Chen, J., Hsu, J. L., Hsu, C. H. and Hua, S. Y. S. Query by Music Segments: An Efficient Approach for Song Retrieval. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2000.

[5] Dannenberg, R. B., Thom, B. and Watson, D. A Machine Learning Approach to Musical Style Recognition. In *Proceedings of the International Computer Music Conference*, 1997.

[6] Doraisamy, S. and Ruger, S. M. An Approach Towards a Polyphonic Music Retrieval System. In *Proceedings of the International Symposium on Music Information Retrieval*, 2001.

[7] Ghias, A., Logan, J., Chamberlin, D. and Smith, B. C. Query by Humming: Musical Information Retrieval in an Audio Database. In *Proceedings of the ACM International Multimedia Conference*, 1995.

[8] Hu, N. and Dannenberg, R. B. A Comparison of Melodic Database Retrieval Techniques Using Sung Queries. In

Proceedings of the ACM Joint Conference on Digital Libraries, 2002.

[9] Huron, D. and Aarden, B. Cognitive Issues and Approaches in Music Information Retrieval. In *Music Information Retrieval*, edited by S. Downie and D. Byrd, 2002.

[10] Jang, J. S. R., Lee, H. R. and Yeh, C. H. Query by Tapping: A New Paradigm For Content-Based Music Retrieval From Acoustic Input. In *Proceedings of the IEEE Pacific-Rim Conference on Multimedia*, 2001.

[11] Kline, R. L. and Glinert, E. P. Approximate Matching Algorithms for Music Information Retrieval Using Vocal Input. In *Proceedings of the ACM International Multimedia Conference*, 2003.

[12] Kuo, F. F. and Shan, M. K. A Personalized Music Filtering System Based on Melody Style Classification. In *Proceedings of the IEEE International Conference on Data Mining*, 2002.

[13] Lu, L., You, H. and Zhang, H.J. A New Approach to Query by Humming In Music Retrieval. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2001.

[14] McNab, R. J., Smith, L. A., Witten, I. H., Henderson, C. L. and Cunningham, S. J. Towards the Digital Music Library: Tune retrieval from acoustic input. In *Proceedings of the ACM International Conference on Digital Libraries*, 1996.

[15] Pickens, J. and Crawford, T. Harmonic Models for Polyphonic Music Retrieval. In *Proceedings of the ACM International Conference on Information and Knowledge Management*, 2002.

[16] Shalev-Shwartz, S., Dubnov, S., Friedman, N. and Singer, Y. Robust Temporal and Spectral Modeling for Query by Melody. In *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2002.

[17] Shan, M. K. and Kuo, F. F. Music Style Mining and Classification by Melody. *IEICE Transactions on Information and Systems,* E86-D(4), 2003.

[18] Tseng, Y. H. Content-Based Retrieval for Music Collections. In *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1999.

[19] Tzanetakis, G., Essl, G. and Cook, P. Automatic Musical Genre Classification of Audio Signals. In *Proceedings of the International Symposium on Music Information Retrieval*, 2001.