Music Classification Using Significant Repeating Patterns

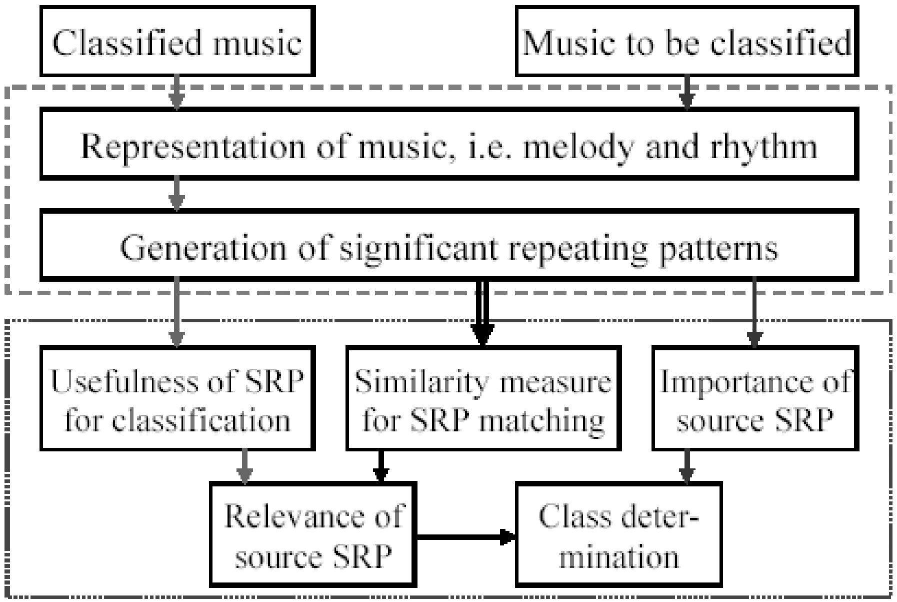
Chang-Rong Lin, Arbee L.P. Chen
Department of Computer Science
National Tsing Hua University, Taiwan, R.O.C

Outline

- Introduction
- Feature Extraction
  - Representations
  - Significant Repeating Patterns (SRP)
- SRP-based Classification
- Experiment Results
- Conclusion

Introduction

- As the amount of music data increases, classification of music data has become an important issue.
- In this paper, we find useful information for classification from the symbolic representations of music data.
- The flowchart of our approach:

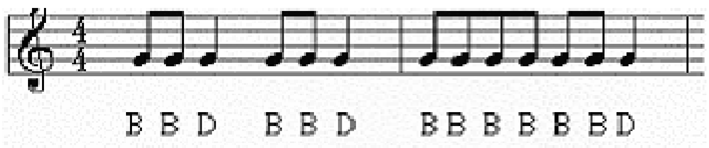


Representations of Music

- Rhythmic sequence

Symbol	Duration	Symbol	Duration	Symbol	Duration
A	(0,1/4]	B	(1/4,2/4]	C	(2/4,3/4]
D	(3/4,4/4]	E	(4/4,5/4]	F	(5/4,6/4]
G	(6/4,7/4]	H	(7/4,8/4]	I	Above 2 beat

The set of beat symbols



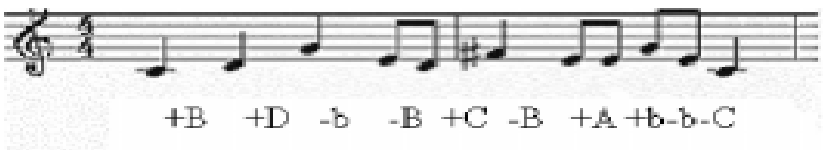
Example of Rhythmic sequence

Representations of Music

- Melody sequence

Symbol	Pitch interval	Symbol	Pitch interval	Symbol	Pitch interval	Symbol	Pitch interval
A	0	B	2	C	4	D	5
E	7	F	9	G	11	H	Other
a	1	b	3	d	6	e	8
f	10	+	Up	-	Down		

The set of pitch symbols



Example of melody sequence

Generation of Significant Repeating Patterns

- Definition of Significant Repeating Patterns (SRP): a consecutive sequence appears frequently in the rhythmic or melodic sequence of a music piece and satisfies the following constraints.
  - Maximum length
    - Reducing duplicate information and the extra costs for pattern discovery
  - Minimum length
    - Alleviating the unnecessary loads due to a large amount of short sequences
  - Minimum frequency
    - The more frequency a sequence has in the music, the more representative it will be.

$Sup(x,m)=\frac{F_{x,m}}{\sum_{\forall SRP\in m} F_{SRP,m}}$

Usefulness of SRP for Classification

- Due to the various lengths of different music, the SRP with a high frequency in one music piece is not necessarily more important than the one with a low frequency in the other, we define *support as*

$$Sup(x,m)=\frac{F_{x,m}}{\sum_{\forall SRP\in m} F_{SRP,m}}$$

$F_{x,m}$ denote the frequency of the SRP *x* for the music piece *m*

- Moreover, for *SRP x* in class *C,* we sum up its support in every music piece belonging to *C* to compute its importance with respect to *C,* which is called the *aggregate support:*

$$ASup(x,C)=\sum_{\forall music\in C} Sup(x,music)$$

Example

Music	Class	SRP (Frequency)	SRP (Support)	Aggregate Support
A	ONE	I(4),II(2), IV(3)	I(0.45),II(0.22), IV(0.33)	I(0.95),II(0.22) III(0.5), IV(0.33)
B	ONE	I(4),III(4)	I(0.5),III(0.5)	
C	TWO	I(2),V(3)	I(0.4),V(0.6)	I(0.4),V(1) VI(0.6)
D	TWO	V (2),VI (3)	V(0.4),VI(0.6)	

Usefulness of SRP for Classification (cont.)

- Owing to the various numbers of music data in different classes, the *SRP* with a high aggregate support in one class is no necessarily more important than the one with a low aggregate support in the other. we further normalize the aggregate support of *SRP x* in class *C* to compute the *normalized support* :

$$NSup(x,C) = \frac{ASup(x,C) - Min(C) + 1}{Max(C) - Min(C) + 1}$$

Min(C): minimum aggregate supports of the *SRP*'s in *C*
Max(C) maximum aggregate supports of the *SRP*'s in *C*

Example

Music	Class	SRP (Frequency)	SRP (Support)	Aggregate Support	Normalized Support
A	ONE	I(4),II(2), IV(3)	I(0.45),II(0.22), IV(0.33)	I(0.95),II(0.22) III(0.5), IV(0.33)	I(1),II(0.58) III(0.74) IV(0.64)
B	ONE	I(4),III(4)	I(0.5),III(0.5)		
C	TWO	I(2),V(3)	I(0.4),V(0.6)	I(0.4),V(1) VI(0.6)	I(0.63),V(1) VI(0.75)
D	TWO	V (2),VI (3)	V(0.4),VI(0.6)		

Usefulness of SRP for Classification (cont.)

- We evaluate the usefulness of each SRP for classification based on its normalized supports in different classes, which is called the *pattern weight :*

$$PW(x, C) = \frac{NSup(x, C)}{TS(x)}$$

TS(x): the total support of *SRP x*.

Example

Music	Class	SRP (Frequency)	SRP (Support)	Aggregate Support	Normalized Support	Pattern Weight
A	ONE	I(4),II(2), IV(3)	I(0.45),II(0.22), IV(0.33)	I(0.95),II(0.22) III(0.5), IV(0.33)	I(1),II(0.58) III(0.74) IV(0.64)	I(0.61),II(1) III(1)IV(1)
B	ONE	I(4),III(4)	I(0.5),III(0.5)			
C	TWO	I(2),V(3)	I(0.4),V(0.6)	I(0.4),V(1) VI(0.6)	I(0.63),V(1) VI(0.75)	I(0.39),V(1) VI(1)
D	TWO	V (2),VI (3)	V(0.4),VI(0.6)			

Similarity Measures for SRP Matching

- Adopting the dynamic programming approach to measure the similarity (i.e. the inverse of *edit distance*) between it and each SRP in a class to identify the corresponding target SRP
- Assigning each symbol (i.e. beat symbol or pitch symbol) a numerical value in order that the difference between two distinct symbols can be computed by a simple subtraction

Beat Symbol	A	B	C	D	E	F	G	H
Value	0.15	0.3	0.45	0.6	0.7	0.8	0.9	1.0

The assigned values of beat symbols

Pitch Symbol	A	B	C	D	E	F	G	H
Value	0.1	0.2	0.3	0.4	0.55	0.7	0.85	1.0
Pitch Symbol	a	b	d	e	f			
Value	0.25	0.35	0.6	0.75	0.9			

The assigned values of pitch symbols

Similarity Measures for SRP Matching (cont.)

- Based on the edit distance, the pattern similarity between two *SRP*'s *x* and *y*, is computed as:

$$PS(x,y) = 1 - \frac{\alpha * D(x,y)}{mleng}$$

D(x,y) the edit distance from x to y
mleng is the maximum constraint on sequence length
The value of *PS*(x,y) between 0 and 1

- Given a source *SRP*, we choose the *SRP* with the maximal value of pattern similarity as the target *SRP* for each class. If more than one *SRP* has the maximal value, we choose the one with the maximal value of pattern weight or the longest one.

Similarity Measures for SRP Matching (cont.)

- Evidence
  - Estimating how a source *SRP* is relevant to a class
  - The formula:

$$E(x,C) = PS(x,y) * PW(y,C), \text{ where } y \text{ is the target SRP of } x \text{ in } C$$

  - Example

PS(X, I)	PS(X, II)	PS(X, III)	PS(X, IV)	PS(X, V)	PS(X, VI)
0.6	0.2	0.8	0.55	0.4	0.5
PS(XI, I)	PS(XI, II)	PS(XI, III)	PS(XI, IV)	PS(XI, V)	PS(XI, VI)
0.4	0.6	0.1	0.3	0.5	0.9

Music	Class	Pattern Weight
A	ONE	I(0.61),II(1) III(1)IV(1)
B	ONE	
C	TWO	I(0.39),V(1) VI(1)
D	TWO	

E(X,ONE)=0.8
E(X,TWO)=0.234

Class Determination

- Each source SRP is associated with two kinds of information:
  - The ***evidence*** indicates its relevance to a class.
  - The ***normalized support*** means its importance with respect to the music to be classified.
- We combine them to estimate the possibility that music m belongs to class C, which is called the *classification score*:

$$CS(C \mid m) = \sum_{\forall SRP \in m} E(SRP, C) * NSup(SRP, m)$$

- The music will be assigned to the class with the highest score

Example of class determination

let the frequencies of the two source SRP’s X and XI be 4 and 2, respectively

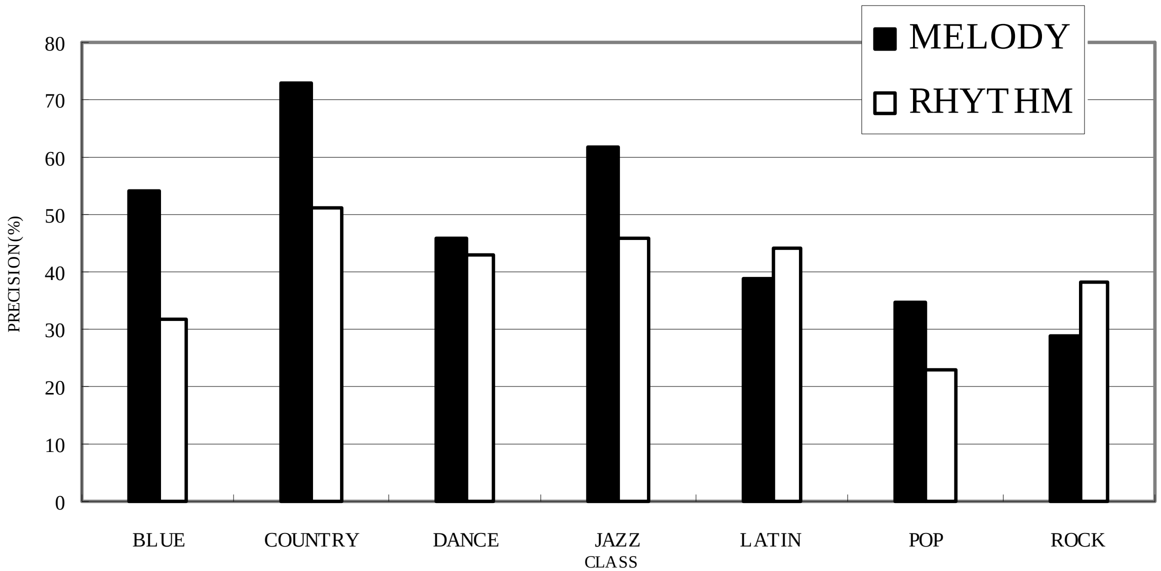
Class	Source SRP (Frequency)	Target SRP	E(x,C)	NSup(x,m)	CS(C|m)
ONE	X(4)	III	0.8	1	1.25
ONE	XI(2)	II	0.6	0.75	
TWO	X(4)	I	0.234	1	0.909
TWO	XI(2)	VI	0.9	0.75	

CS(ONE|m)=0.8*1+0.6*0.75=1.25
CS(TWO|m)=0.234*1+0.9*0.75=0.909

Highest score

Experiment Results

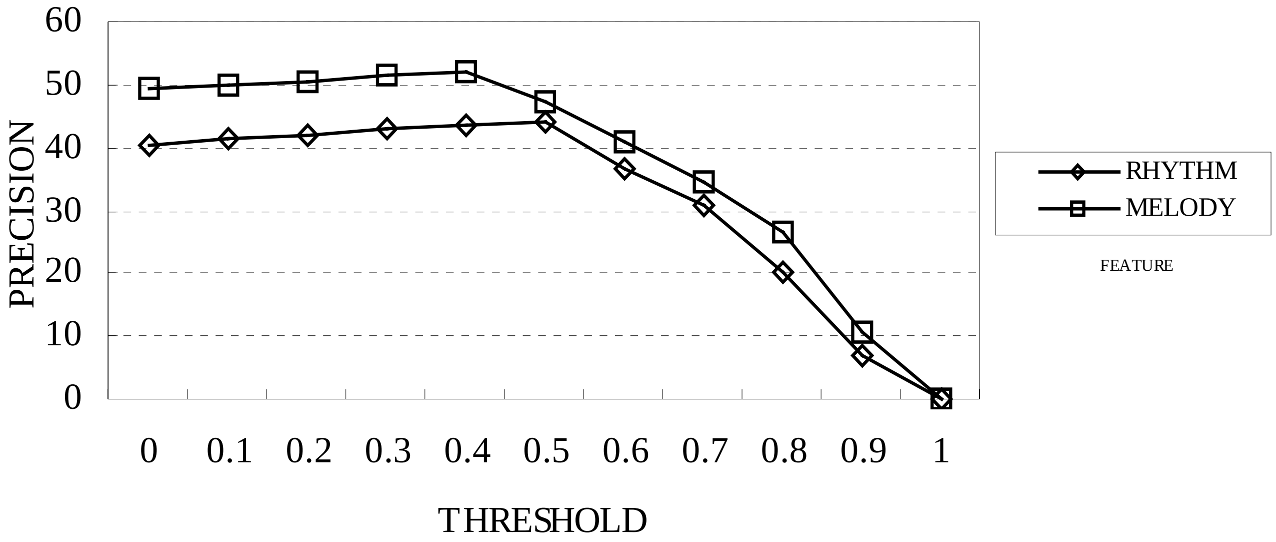
- Impacts of Features



The precision for different features in the seven classes

Experiment Results (cont.)

- Impacts of Similarity Threshold

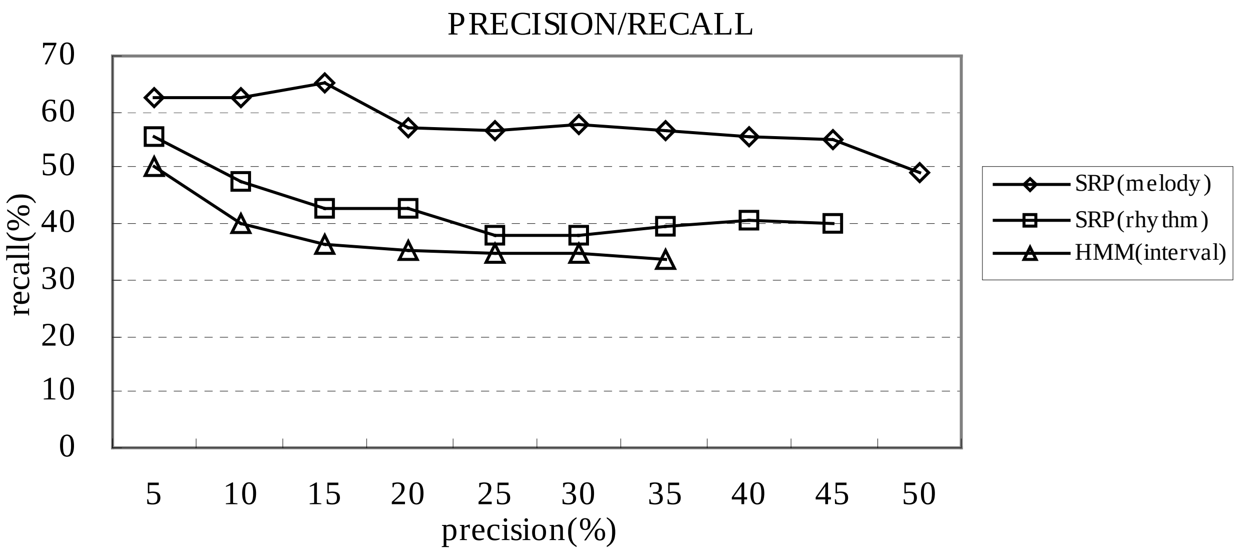


Experiment Results (cont.)

- Comparison with the HMM-based Approach

	SRP (melody)	SRP (rhythm)	HMM
Precision (%)	49.18	40.24	33.70

The comparisons on the average precision



The diagram of precision and recall

Conclusion

- We present a scheme for generating significant repeating patterns.
- A way to estimate the usefulness of SRP for classification is also proposed.
- For the music to be classified, we incorporate human perception and musicology into the similarity measures for SRP matching.
- The experiment results indicate that some classes achieve better precision for a particular feature.
- This approach performs on average better than the HMM-based approach.