Knowledge Discovery over English Songs

Hu Jiabin, Huang Richeng, Wangbo, Zhao Jing Advised by : Prof. Raymond Wong



Introduction

Music plays an important role in people's life, without which the essence of society will fly apart. With developments of digitalized music technologies, process of music composition, compilation, appreciation and sharing have experienced great changes.



1.Find how a sequence of stresses from lyrics is matched with the melody. we call this a pattern.

2.Generate a melody automatically for given lyric according to patterns found in Taks1 3. Find a number of groups where each group has "similar" patterns

Goals

completed



Fig.1: User interface

Features

Freely set the beat, major, speed and base octave for songs
Manually modify specific note according to users' preference
Detailed information shows in the user interface

Input Lyrics

Match each syllable in lyrics with notes according to frequent patterns

Generate the whole melody

Fig.2: Basic flow of how to generate melody

Working Principle

The project mainly consists of five parts

1. Preprocessing

Lyric and melody of midi songs were digitalized with the help of CMU Pronouncing Dictionary and NoteWorthy.

2. Find sub-frequent patterns

Patterns with occurrence frequency larger than specified threshold in every midi song was marked as sub-frequent patterns.

3. Establish frequent pattern pool

After getting sub-frequent patterns, we treated them as "ordinary" and then found "frequent" patterns in the range of the all midi songs to form the frequent pattern pool.

4. Match lyric with melody

Lyrics input by users were matched with certain notes based on frequent pattern pool.

5. Enhance music quality

Additional features like speed, beat, major were added to improve the quality of melody



Fig.4: Procedure of forming frequent pattern pool



		J Weka Explorer	Weka Clusterer Visualize: 22:40
Digital input lyric from user: -3,2,1,-1 2,-1,3,2,-3	-3,2,1,-1, 000 -3,2,1,-1,	Freprocess Classify Cluster Associate Select attributes Visuali Clusterer Choose SimpleHMeans -W 4 -A "weka core EuclideanDistance -R Cluster mode Clusterer out	re first-le put I: Instance_number (Num) I: Cluster Colour: Cluster (Nom) Select Inst Jitter
Matched candidate melody	If not matched, select one of the top <i>n</i> Frequent	Vse training set Supplied text set Set Percentage split K 86 Classes to clusters evaluation	0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
[Melody]:Probability P [2,-1,4,-3]:0.6 [1,-2,3,-1]:0.2 [2,-2,5,-3]:0.03 [2,-1,0,0]:0.17	Pattern from All Frequent Pattern Set F	(Itum) 711 ~ (Itum) 711 ~ (Itum) 711 ~ (Itum) 709 709 709 710 710 711 (Itum) 709 710 711 710 712 710 711 710 712 710 713 710 714 710 715 710 710 710 711 710 712 710 713 710 714 710 715 710 716 710 717 710 718 710 719 710 710 710 710 710	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
hoose a melody sequence andomly, based on P	T T		and eva
2,-1,4,-3		• • • • • • • • • • • • • • • • • • •	0 355.5 711
Output melody sequence		Status OK Log	cluster0 cluster1 cluster2 cluster3

Fig.5: Procedure of matching lyric with melody

Fig.6: K-mean clustering by Weka

Implementation

In our project, we established music library by retrieving midi files with different styles and digitalized the lyric and melody information of songs in the music library. We used the pitch of melodies and stress of English words as features for data mining and generated a pool of frequent patterns. With these patterns, we generated new melodies given lyrics and also used Jaccard coefficient to measure similarity between songs for further clustering by Weka.

Evaluation & Testing

- Periodic debugging and testing were conducted to ensure correctness of programs.
- A survey was carried out to evaluate final effectiveness of data mining.



Survey on Composer

1.What do you think of our webpage design?

	Excellent	Good	Moderate	Acceptable	Poor
Aesthetics	0	0	0	0	0
Operability	0	0	0	0	0
Functionality	0	0	0	0	0
	Aesthetics Operability Functionality	Excellent Aesthetics Operability Panctionality O	Excellent Good Aesthetics O O Operability O O Functionality O O	Excellent Good Moderate Aesthetics O O O Operability O O O Functionality O O O	Excellent Good Moderate Acceptable Aesthetics O O O O Operability O O O O Functionality O O O O

2. How do you evaluate the quality of sample music?

	Excellent	Good	Moderate	Acceptable	Poor
Rhythm	0	0	0	0	0
Melody	0	0	0	0	0
Expression	0	0	0	0	0

3.Will you recommend our webpage to your friends?

O Yes

4.How do you rate our overall project?

1 2 3 4 5 6 7 8 9 10

0 0 0 0 0 0 0 0 0 0

5.Any more improvement for our webpage?

「現文」 切勿通过 Google 表単提交密码。

L Google

Result

In the project, we generated a pattern pool and came up with a system generating melodies given lyrics using frequent pattern matching. Our webbased user interface allows users to input and obtain returned generated melodies