



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

Goals completed

3. Find a number of groups where each group has "similar" patterns

Composer

[Main](#) [About us](#)

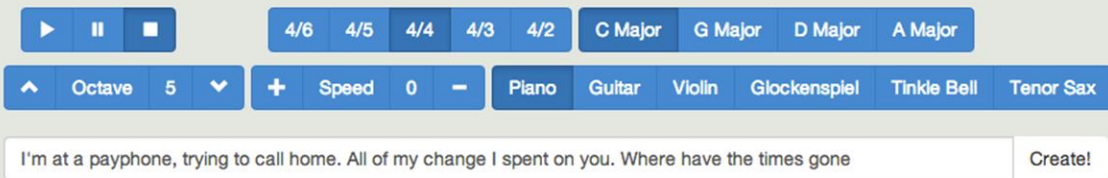


Fig.1: User interface

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

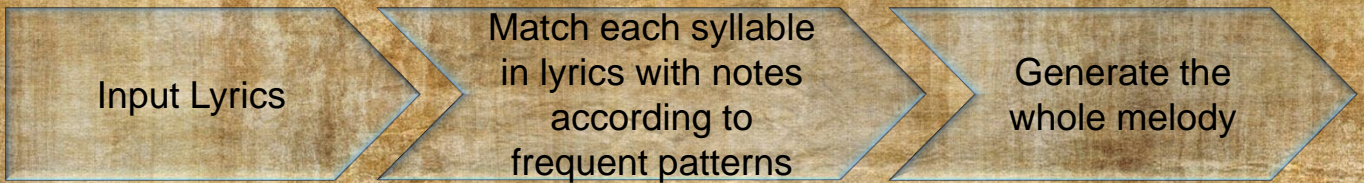


Fig.2: Basic flow of how to generate melody

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.3: NoteWorthy

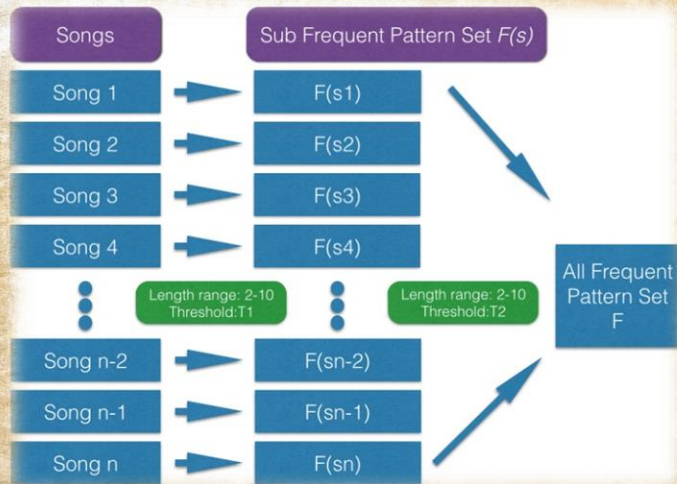


Fig.4: Procedure of forming frequent pattern pool



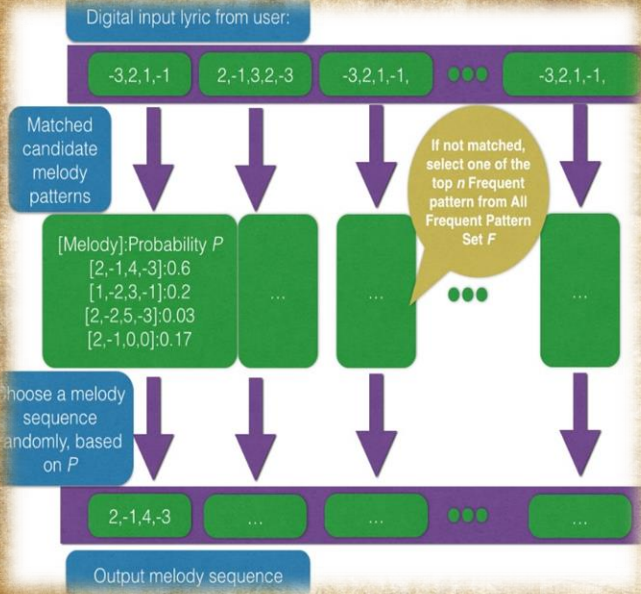


Fig.5: Procedure of matching lyric with melody

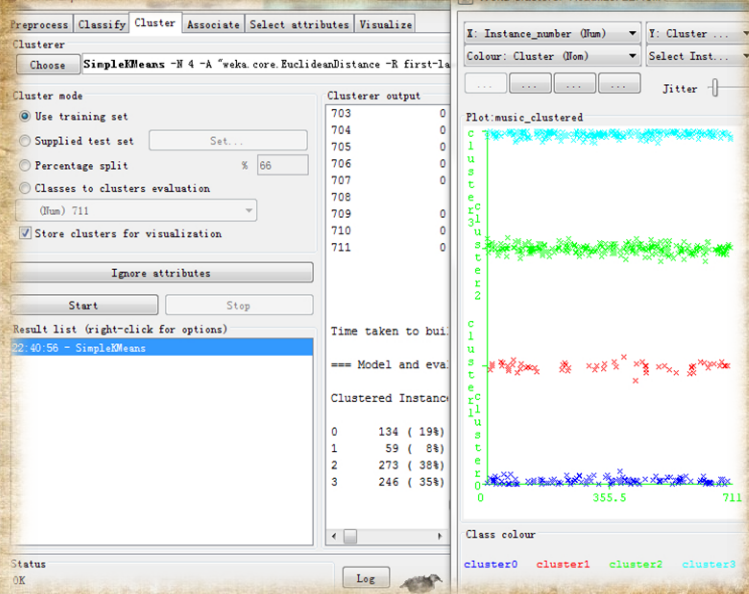


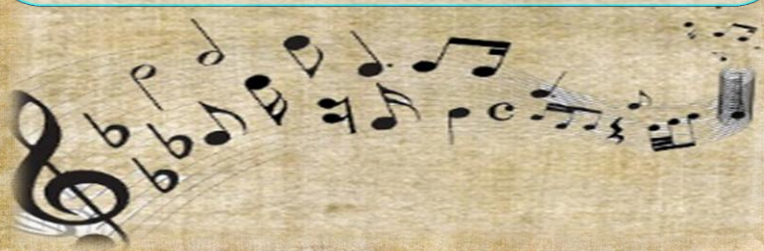
Fig.6: K-mean clustering by Weka

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.



Result

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

Survey on Composer

1.What do you think of our webpage design?

| | Excellent | Good | Moderate | Acceptable | Poor |
|---------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Aesthetics | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Operability | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Functionality | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

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

| | Excellent | Good | Moderate | Acceptable | Poor |
|------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Rhythm | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Melody | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Expression | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

3.Will you recommend our webpage to your friends?

Yes
 No

4.How do you rate our overall project?

1 2 3 4 5 6 7 8 9 10

5.Any more Improvement for our webpage?

请勿通过 Google 表单提交密码。

Fig.7: Survey