



Music rhythm customized mobile application based on information extraction

Yining Li, Wei Hu and Yonghao Wang

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

September 6, 2019

Music rhythm customized mobile application based on information extraction

Yining Li¹, Wei Hu² and Yonghao Wang³

¹ College of Computer Science and Technology, Wuhan University of Science and Technology, China

² Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan, China

³ Digital Media Lab, Birmingham City University, UK
wustlyn@gmail.com
huwei@wust.edu.cn
Yonghao.wang@bcu.ac.uk

Abstract. Information extraction technology to be able to measure, store, collect all kinds of information, especially the direct access to important information, which is based on mobile applications and more convenient for the information gathering process, user and information feedback, greatly reduce the cost of information technology, makes the implementation of large-scale information extraction technology possible. As for this paper, first of all, mainly introduces the basic theory of music rhythm customization mobile application; Secondly, the development and implementation of this application are introduced; Finally, it summarizes and anticipates the future development trend of music rhythm customization technology. After implementation, users only need to import music or video that they want to modify, select the corresponding style or double speed, and then get relevant audio results through system software processing. And make music rhythm customization easy to operate, which remove a lot of irrelevant operations, so that users do not need to know the relevant professional knowledge can be processed.

Keywords: Rhythm tracking, Audio processing, Information extraction, A mobile application.

1 Introduction

The rhythm tracking mainly uses information extraction technology. This technology was first used in surveying and mapping to extract useful information for users in remote sensing images [1]. The earliest possible application of this technology to the field of music was the feature extraction of phrase-based music information retrieval published by Japanese scholars in 1999.

In this paper, the audio data is converted into an audio stream, and the audio data is transformed into a data frame by the rhythm extraction algorithm by using the frame

changeable characteristics of the audio data stream, and the audio is clipped by using the selected double speed time stretching algorithm, and then the template is used. Process audio frames, create/edit your own audio, save it in the sandbox or share it with your friends. Developers only need to add their own customized templates, users only need to import the audio they need to modify, adjust the paragraphs that need to be modified, and select the corresponding speed and template to get their own music. In the end, this paper designed a rhythm customization application centered on information extraction, which realized a simple and clear audio editing method.

Rhythm tracking technology is essentially a technique that converts the music data imported by the user into an array of rhythms and edits the audio array accordingly [2]. The music rhythm customization technology based on information extraction is used for mobile application development. With the rhythm tracking algorithm and time stretching algorithm of ffmpeg audio and video framework on audio, the effect of customized music is achieved: 1) It is easy to operate, removes a lot of unrelated operations, can be processed without the user knowing the relevant professional knowledge; 2) rich rhythm templates are available; 3) free time stretching, software supports the start and end time interception and speed editing of imported audio.

2 Related Work

2.1 rhythm tracking algorithm

There is no doubt about the importance of data. With the growing popularity of big data, the importance of data information extraction, data analysis and processing is also becoming increasingly prominent. At present, in addition to theoretical research on big data, there are also ongoing technical research abroad, with the focus on data application and data engineering. In China, the development of big data is relatively young. In the 13th five-year plan released by the Chinese government on November 3, 2015, the term "big data strategy" was put forward for the first time [3]. At present, the research of big data in China mainly lies in distributed storage and cloud computing. Only a few enterprises can use the integrated process of data acquisition, data analysis and data processing.

According to the current situation, information extraction system and data control system have been closely combined and developed in two directions: one is distributed information extraction and data control system, and the other is distributed information extraction and data control system. Information extraction technology is the data source of big data technology, and big data technology is the support of the current hot artificial intelligence technology. Therefore, the information extraction technology in the future will still develop at a very fast speed and towards the direction of extracting massive data.

In this paper, an iOS platform mobile application named LickMixeR-N was designed, which realized the rhythm tracking algorithm and time stretching algorithm based on ffmpeg audio and video framework, and used it in audio to realize the effect of changing music and video imported by users and realizing customized music. In this application, users only need to import the music or video they want to modify,

select the corresponding style or double speed, and then through the system software processing to get relevant audio results.

3 Proposed Model

3.1 rhythm tracking algorithm

Subsequent paragraphs, however, are indented. The main rhythm tracking algorithm consists of three steps: intermediate input representation, general state, and contextual state. The middle input representation, also known as rhythm detection function, converts the input audio signal into an array of audio frames, serving as an intermediate signal between the input audio rhythm and the output audio rhythm. The tracking formula is shown in formula (1) :

$$\Gamma(m) = \sum_{k=1}^K |S_k(m) - \hat{S}_k(m)|^2. \quad (1)$$

In general, the main task is to detect the rhythm periodicity and the rhythm alignment without knowing the audio input [4]. The main task of context-dependent state is to merge and extract context-dependent information (including rhythm, time signature and past rhythm location) from audio input into relevant parameters [5]. In order to achieve the effect of emphasizing significant rhythm and discarding insignificant rhythm, the adaptive moving average range of a frame is calculated as shown in formula (2) :

$$\bar{\Gamma}_i(m) = \text{mean}\{\Gamma_i(q)\} \quad m - \frac{Q}{2} \leq q \leq m + \frac{Q}{2} \quad (2)$$

Rhythm tracks are cut into the frame, query beat alignment, contextual, cycle a few characteristics, and the general state of running in the form of no memory, extracted by repeating rhythms cycle and alignment, the disadvantages of this method lies in the context, makes the final results do not have continuity, and the strong consistency we hear music is a big difference [6]. Single use general state can generate many unnecessary mistakes, therefore, need to introduce the contextual state is used to solve and separately to solve the problem of continuous audio output.

3.2 Time stretch and audio spectrum display

Time stretch means to stretch the length of input audio [7]. The source audio is cropped by selecting the start time and end time. The audio spectrum display uses the fast Fourier transform algorithm to transform the input audio into the corresponding spectrum array [8], which is finally displayed on the interface.

Time stretch can call the function in rubber band Audio, input the Audio rhythm array and corresponding doubling speed, select the corresponding conversion mode,

and then get the result rhythm array[9];Then the result array is transformed into a spectrum image by fast Fourier transform and displayed on the interface in real time.You can use the time stretch function to complete the audio processing[10].

4 Experiment

4.1 Audio Production

First of all, since this app is a pure iOS platform audio mobile app, what needs to be considered is how to import local audio[11]. Apple has made the API of audio selection for developers in its own framework. Just import the relevant framework and you can easily use MPMediaItem to store audio files of iOS.

Second, this article describes an audio editing application, so the second step is to process the audio just obtained, including the intermediate input layer representation, general state processing, and context-dependent state mentioned above in rhythm trackingTheory, this part is mainly the use of rhythm tracking algorithm to achieve;

Then through the user's input to the audio file time and double speed editing, this part mainly USES the time stretch technology to achieve;

Finally, after editing audio in local or share out, for users to upload audio data security, did not use the cloud database, but instead USES application sandbox to save the audio data for the user, the user can see myself in the historical view of all the audio data, and view or delete it[10];If users want to export/share their results, they can upload audio files from the sandbox to DropBox, and then further share/export.The workflow of the whole application is shown in figure 3.1:

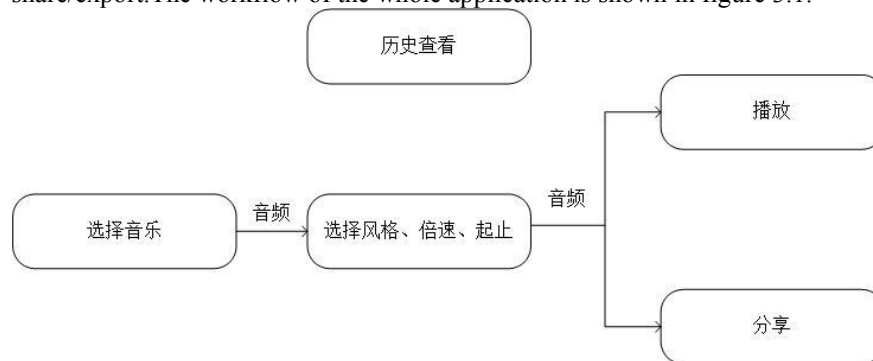


Fig. 1. A figure caption is always placed below the illustration. Short captions are centered, while long ones are justified. The macro button chooses the correct format automatically.

Can be seen from the diagram, the design of the audio file operations is the main process of the one-way, and users can at any time to back up one step operation, but not finished step, as long as the user will not be able to take the next step operation, so the application interface card operation, background using the album cover (the url of the picture in the MEMediaItem albumImage attribute), will be applied to each function into one card[12]. Then according to the steps to use ScrollView paging was

carried out on the card, and then set the card swiping to step on the unfinished disable mode[13].

4.2 Development Of Implementation

In this paper, the development environment of music rhythm customized mobile application based on information extraction is taken as an example under the MacOS operating system [14]. It uses CocoaPods third-party library introduction tool to import DropBox library, directly introduces FFMPEG audio and video processing framework and Fourier transform for operation, and uses iOS development tool Xcode for development [15]. FFMPEG is a very powerful audio and video processing framework developed by iOS. Its main functions include recording, converting digital audio, video, and converting them into data stream, audio and video encoding/decoding, etc [16].

The main design idea is to introduce FFMPEG framework from the outside, configure CocoaPods, import algorithm function and related files into the framework using ovocaine framework and other fast Fourier transform framework to play, and finally use Xib file based on Interface Builder tool to describe part of the view in a graphical way [17].

5 Conclusion

As a symbol of the inheritance of human civilization and a weathervane of times culture, music is the best mark of an era, and editing music only belongs to oneself has become the interest of many people [18]. After the rise of the contemporary Internet, music entertainment has become the just need of people's spiritual entertainment, and editing music only belongs to themselves has become the interest of many people. The music rhythm customization mobile application described in this article based on information extraction makes music editing no longer remote for non-professionals, and realizes the customization of music rhythm. The music rhythm customization mobile application described in this article based on information extraction makes music editing no longer remote for non-professionals, and realizes the customization of music rhythm [19]. This paper presents a music customized mobile application based on information extraction. By applying rhythm extraction and rhythm tracking algorithm to the audio submitted by users, users can turn their little inspirations into audio output that actually conforms to the definition of music template with simple operation. Due to the customization of uploaded audio, users can generate their own music [20]. In short, the music customized mobile application based on information extraction implemented in this paper does not require high design and development costs, and is of great practical value due to its simple operation, simple page and stable function.

References

1. Yu zhanjie, F.: The positioning, reasons and significance of the relationship between rites and music in early Confucianism. Film review and introduction(18), 106-109(2010).
2. Ravi N D, F.: Bhalke D G, S.: Musical Instrument Information retrieval using Neural Network (2016).

3. Wang yue, F.: Concept and practice of data journalism in the context of big data . *Modern media* 23(6), 16-17 (2015).
4. Fu huazheng, F., Chen chong, S., Xiang yong, T., et al: Research and implementation of key technologies for distributed big data acquisition. *Guangdong communications technology* 35(10),7-10 (2015).
5. Davies M E P, F ., Plumbley M D, S.: Context-Dependent Beat Tracking of Musical Audio. *IEEE Transactions on Audio, Speech and Language Processing* 15(3), 009-1020(2007).
6. Chu W, Champagne B, F .: Further studies of a FFT-based auditory spectrum with application in audio classification, *International Conference on Signal Processing* 1-3 (2008).
7. Degara N, F., Rua E A, S., Pena A ,T., et al: Reliability-Informed Beat Tracking of Musical Signals[J]. *IEEE Transactions on Audio, Speech and Language Processing*20(1), 290-301 (2012).
8. Mohapatra B N, F., Mohapatra R K, S.: FFT and sparse FFT techniques and applications, *Fourteenth International Conference on Wireless & Optical Communications Networks. IEEE*, 1-5 (2017).
9. Zhan Y, F., Yuan X, S.: Audio post-processing detection and identification based on audio features, *International Conference on Wavelet Analysis & Pattern Recognition. IEEE*, 3-7(2017).
10. Greamo C, F., Ghosh A, S.: Sandboxing and Virtualization: Modern Tools for Combating Malware, *IEEE Security & Privacy*, 9(2), 79-82 (2011).
11. Pang bo, F.: Development trend of interface design from flat style , *Decoration* (4),127-128 (2014).
12. Roig C, F., Tardón L J, S., Barbancho I, T., et al: Automatic melody composition based on a probabilistic model of music style and harmonic rules[J]. *Knowledge-Based Systems* 71,419-434 (2014).
13. Marchand S, F.: Fourier-based Methods for the Spectral Analysis of Musical Sounds, *Signal Processing Conference. IEEE*, 1-5(2014).
14. Li ping, F., Zou zhenyu, S.: Loop of the scroll view class (UIScrollView) in apple iOS and algorithm of dynamic image loading , *Computer & telecom*, (10), 54-55 (2011).
15. Liu chenji, F., Zhou baogang, S., Guo shuqin. T.: Load optimization of large quantity data based on UITableView in iOS [J]. *Journal of hangzhou university of electronic science and technology*, (4), 46-49 (2013).
16. Ma Z, F.: Digital rights management: Model, technology and application, *China Communications* 14(6), 156-167 (2017).
17. Kim B, F., Pardo B. S.: Speeding Learning of Personalized Audio Equalization, *International Conference on Machine Learning & Applications*, 3-6 (2015).
18. Rong F, f.: Audio Classification Method Based on Machine Learning, *International Conference on Intelligent Transportation. IEEE Computer Society*, 3-5(2016).
19. Liu lixia, f., Bian jinsong, S., Zhang li T., et al: Implementation of audio and video synchronization based on FFMPEG decoding , *Computer engineering and design* 34(6), 2087-2092 (2013).
20. Akram F, F ., Garcia MA, S.: Puig D Active contours driven by local and global fitted image models for image segmentation robust to intensity inhomogeneity. *Plos One* 12(4), 1 – 32 (2017).