

Research On Classification Of Heart Sound Signal Based On Multi Scale Features

Zhang Lejuan¹, Cui Shiyao¹, Zhang Lu¹, Li Niqiang¹

¹School of Information Science and Engineering, University of Jinan, Shandong Jinan, 250022, China

^aEmail: Lejuanzhang33@163.com

Keywords: Heart Sound Signal; Multi Scale Features; Classification

Abstract. Since the near duplicated video detection in this study has obvious relevance, namely testing two videos are the same or nearly the same, thus the method in this article does not focus on the structural analysis of the video, but on two main points: 1) effective elimination of redundant video frames to improve the detection efficiency; 2) maintaining the granularity of the extracted key frames to make sure the accuracy of test results. Based on these two objectives, the thesis proposes an automatic dual-threshold method to eliminate redundant video frame, divides the video into several video segments, and takes the methods of adopting fixed extracted frame rate on each video segments to extract three key frames to represent the video segment on each video. This method can effectively eliminate the redundant frames of video, but also ensure extracting the granularity of key frame, which achieves good practice effect

Introduction

The important inner feature of video is its time characteristic, i.e. the video is composed of successive video frames in the direction of the time, thus content-based video retrieval or near repeated video judgment usually need to be compared the similarities of the video frames set to judge the similarity between two videos, although, in principle, content-based near duplicated video detection is basically the same with the content-based video retrieval. However, in practical applications, there is a big difference in terms of implementation and existence. Content-based video retrieval pays more attention to the overall similarity of the query video and the goal video, that is, more of a "one to one" relationship, while for near duplicated video detection, its similarity retrieval may face more matches in the form. These matching forms can be summarized as "one-one", "one-many" and "many-many"[2]. These three forms are described in Figure 1. Because the existing forms of near repeated video are entirely dependent on the specific application of the target video. If the near duplicated video is operated by complex editing and modification, its detecting work will be a typical matching problem of video sub-sequence. In the complex in the complex near duplicated video detection, the length of the target sub-sequence, the existing location and the existing frequency are unknown (the form of "many to many" showing in Figure 1 (c)); this form of matching task of video subsequence is much more complex than the ordinary video retrieval tasks.

A Graph-based Matching Algorithm on Sub-sequence of near Duplicated Video

In order to facilitate the following discussion, this paper defines the related concepts as follows:

$Q = \{q_1, \dots, q_{|Q|}\}$ and $R = \{r_1, \dots, r_{|R|}\}$ are respectively represent the key frame collection of query video and reference video; q_i represents the i frame of the query video, and r_j means the j frame of the reference video; $|Q|$ is the number of video frames and $|R|$ represents the number of reference video frames.

$G = (N, E)$ means the matching results graph; N means the matching nodes collection; E means the set of edges.

$\mathcal{N} = \{M_{i,j} | 1 \leq i \leq Q, 1 \leq j \leq R\}$. Node $M_{i,j}$ in the nodes collection represents the existing match which can satisfy specified constraints between the query video frames q_i and the reference video frame r_j . The most common used constraint is to result the similarity which is larger than a specified threshold τ between the query video frames q_i and the reference video frame r_j . In addition, this also can be specified by $K-NN$, that is, the inquiry frame q_i returns to previous k nearest reference video frame.

$E = \{<M_{i,j}, M_{l,m}>\}$. $<M_{i,j}, M_{l,m}>$ means that there is a line between the node $M_{i,j}$ and the node $M_{l,m}$.

$sim(q_i, r_j)$ indicates the similarity between the query video frames q_i and the reference video frame r_j .

$P = \{p_1, \dots, p_n\}$ shows the set of all the matching paths in the graph G .

$\bar{Q} = \{q_m, \dots, q_n\}, 1 \leq m < n \leq Q$ and $\bar{R} = \{r_m, \dots, r_n\}, 1 \leq m < n \leq R$ are respectively indicate the sub-sequence of query video and the sub-sequence of the reference video.

$sim(p_j)$ means the similarity between the sub-sequences of two corresponding video in the matching route.

$S_Q = \{a_1, \dots, a_L\}$ and $S_R = \{b_1, \dots, b_L\}$ are respectively represent the frame ID set corresponding to query video and the frame ID set corresponding to reference video in the matching route of the video subsequence, which means that in the matching path, the a_l frame of the query video and the b_l frame of the reference video are one-to-one correspondence; L represents the length of the matching path.

In this thesis, the author converts the matching results of the video in the space information into the matching results graph, and then changes the location problem of near duplicated video into the problem of searching the key route in the matching results graph. The specific flow is shown in graph 2. It's main steps include: 1) extraction of video key frames; 2) features matching based on key frames; 3) building the matching results graph based on matching results; 4) detection of matching route; 5) filtering matching route; 6) location of near duplicated sub-sequence.

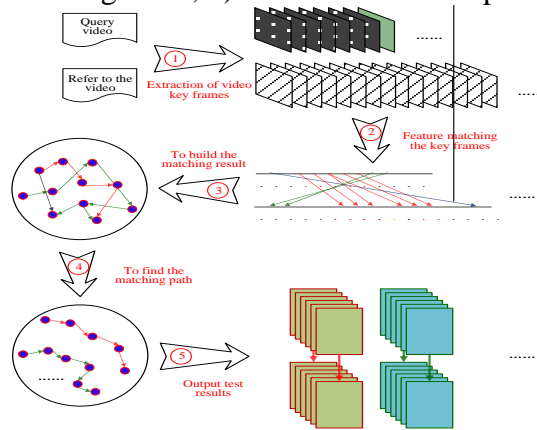


Figure1. Flowchart of video subsequence matching based on graph

The flowchart of video subsequence matching based on graph in Figure 2, and the key frame extraction is the first step based on the near duplicate video detection of the contents. The accuracy and granularity of extracting the video key frame are largely determines the quality of detecting the near duplicated video. Video is consisting of a sequence of frames with the timing characteristics, and when the human visual characteristics are taken into account, the similarity of adjacent video frames will be high. This makes the visual information in the successive video frames has a great redundancy in the time direction, and if during the detecting the near repeated video and extracting the features to each frame, the extra unnecessary calculations will be brought out. So the representative video key frames need to be extracted to reduce redundancy of visual information. Thus, when the near duplicated video is detected, there only needs to match the key frame sequence, thus the cost of accounting can be greatly reduced in the matching process.

According to the above definition, the matching task of video subsequence also can be defined as

follows:

Video subsequence matching: in a predefined score function, the most similar part corresponding to the reference video sequence and the query video sequence is searched, that is, two video sub-sequences \bar{Q} and \bar{R} are searched to make highest similarity. In response to the abrupt and gradual of the video visual content, this paper adopts the dual-threshold method similar to the method proposed in literature [6]. There are two characteristics of this method as following: 1) usage of a dual-threshold method; one threshold T_h is adopted to detect the abrupt of frame visual information, and the other threshold T_l is adopted to detect the gradual of frame visual information. 2) The size of the two thresholds is not pre-determined, but according to the change of visual information of the video the size of threshold is automatically determined. Threshold T_h and threshold T_l are defined as follows: $T_h = \mu + \alpha\sigma$, $T_l = \beta T_h$. μ and σ are respectively mean and standard deviation of the entire video frame based on color features. According to the studies in literature [24, 6], α is recommended to set between 5 and 6. According to the studies in literature [23], β is usually chosen from 0.1 to 0.5.

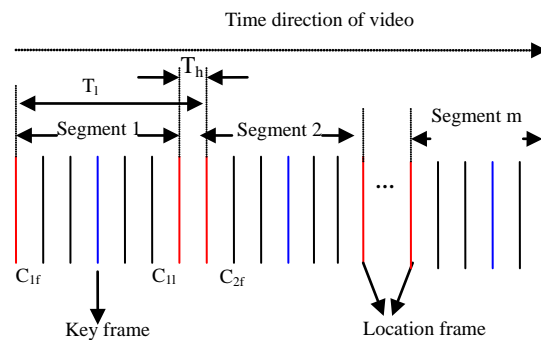


Figure2. Elimination of the video redundancy frame and extraction of the key frame on f adaptive dual-threshold

Experimental Results

If there is an edge between two nodes in the matching figure, there are two conditions that must be simultaneously met as following:

- (1) The two nodes must be satisfied the consistency in the time direction;
- (2) The time jump degree of two nodes is $\Delta t < \tau$ (τ is threshold of time jump degree)

The condition (1) represents the time direction of the query video represented by two nodes is consistent with the time direction of the target video graph, which may seems reasonable to do so, because the video sequence is a time s sequence and the time direction of the coping video and the copied video is the same. The direction of increasing time is directed edges between nodes direction. The condition (2) represents the jumps of the matching results represented by two nodes in the time direction can not exceed a certain threshold, otherwise there is no correlation between the two matching results. According to the above methods and conditions, the matching results in figure 5 can be converted into a matching figure, and there is obviously a directed acyclic graph in the matching results figure. The matching results of the eight rows in Figure 5 can be converted into the matching results shown in Figure 6. The author also had a comparative experiment to obtain the most optimal time jump threshold as shown in Figure 7; T1-T10 in the figure are represent 10 different types of copied video (For details, see [1]); the experimental data shows that a relatively appropriate time jump threshold should be between 10 to 29 seconds. On the other hand, the author also makes the statistics to the time length of the obtained video clips of the adaptive dual threshold method as shown in Figure 8. The time length of video segments shown by data in the figure mostly concentrates between 10 and 30 seconds, which also verified the reasonability of choosing threshold of time jump degree between 10 and 30 seconds.

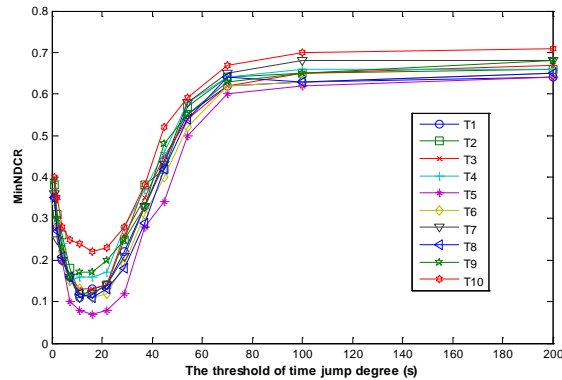


Figure3: selection of threshold of time jump degree

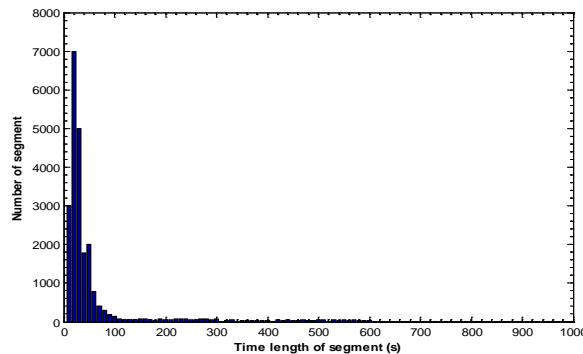


Figure 4: distribution of time length of segments after the video is divided by adaptive dual threshold method

Conclusions

Because there are lots of uncertain factors existing in process of near duplicated video detection, such as whether exist near duplicated video in the target vide, how long the near duplicated video segments are and what position the near duplicated video locates, which makes the detection of near duplicated is harder compared with the usual video retrieval. Currently, there are two main drawbacks of the subsequence matching algorithm for near duplicated video detection: Firstly, a similar threshold is needed to determine in the matching process to decide the returned results, while for different videos, it is really difficult to determine a common threshold. Secondly, it is cost a lot to detect the copied video with all possible lengths and possible positions by using exhaustive. So this thesis proposes a matching algorithm of video sub-sequence based on graph. This method skillfully converts video sequence matching results into the matching results graph, and then converts the video copy detection into a problem of finding the longest path in the matching results figure. There are several distinct advantages of this method based on graph: (1) a graph-based approach can find out the most optimal matching sequence in a number of clutter, which effectively eliminate some noises brought by some false "high similarity" matches, so it can make up inefficient defects of describing the underlying characterization. (2) Because it makes full consideration and use of the timing characteristics of the video sequences, the positioning accuracy of near duplicated video is really high. (3) Graph-based sequence matching method can automatically detect the plurality of discrete paths, so it can locate in the batch several near duplicated video in these two videos. (4) Compared with the exhaustive testing, this method has obvious advantage on the detection speed.

Acknowledgement

In this paper, the research was sponsored by the School Foundation: XKY1131

References

- [1]Wang Lina, Gao Hanjn, Liuwei, Peng yang. Detection and management of virtual machine monitor. Research and development process of Computer, 2011, pp: 1534-1541.
- [2]Yinghua Xue, Hongpeng Liu, Intelligent Storage and Retrieval Systems Based on RFID and Vision in Automated Warehouse. Journal of Networks, Vol. 7, No. 2 (2012), pp: 365-369
- [3]Huan Zhao, Kai Zhao, He Liu, Fei Yu, Improved MFCC Feature Extraction Combining Symmetric ICA Algorithm for Robust Speech Recognition, Journal of multimedia, Vol. 7, No. 1, 2012. pp: 74-81
- [4]Y. Geng, J. Chen, K. Pahlavan, Motion detection using RF signals for the first responder in emergency operations: A PHASER project, 2013 IEEE 24nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), London,Britain Sep. 2013
- [5]Y. Geng, J. He, K. Pahlavan, Modeling the Effect of Human Body on TOA Based Indoor Human Tracking[J], International Journal of Wireless Information Networks 20(4), 306-317