International Journal for Modern Trends in Science and Technology
Volume: 03, Issue No: 05, May 2017
ISSN: 2455-3778
http://www.ijmtst.com

Text Feature Selection and Extraction over Video Segmentation and Retrieval

K.Ramya[1] | M.Siva Sundara Vinayagamoorthy[2]

[1,2]Department of CSE, AVC College of Engineering, Mayiladuthurai, Tamilnadu, India.

To Cite this Article
K.Ramya and M.Siva Sundara Vinayagamoorthy, "Text Feature Selection and Extraction over Video Segmentation and Retrieval", *International Journal for Modern Trends in Science and Technology*, Vol. 03, Issue 05, May 2017, pp. 275-278.

ABSTRACT

Digital video has become a popular storage and exchange medium due to the rapid development in recording technology, improved video compression techniques and high-speed networks in the last few years. As a result, there has been a huge increase in the amount of multimedia data on the Web. Therefore, for a user it is nearly impossible to find desired videos without a search function within a video archive. Even when the user has found related video data, it is still difficult. Most of the time, the user was unable to judge them whether the video is useful or not. At the same time, it is also an opportunity to obtain the information about major events, after glancing at the search results through web video mining. To address this need, web video event mining approaches have been developed. In this paper, a novel framework is proposed that integrates thevisual temporal information and textual distribution informationfor web video event mining.

Keyword- *Co occurrence, multiple correspondence analysis (MCA), near-duplicate key frames (NDK), visual feature trajectory.*

I. INTRODUCTION

Search engines and video sharing websites such as YouTube, Google, make it convenient for the users to access relevant web video. The data show that over 100 h of videos are uploaded to YouTube [1] every minute, and six billion hours of videos are watched each month. These facts demonstrate a new challenge for the users to grasp the major events available from searching video databases. At the same time, it is also an opportunity to help them obtain the information about major events, after glancing at the search results through web video mining. The users have to browse many web videos in the returned list and even watch most of the videos. This is not only time consuming, but also difficult if thousands of web videos are returned by a search engine. The TDT detects the topics and tracks known events in text. TDT is mainly is focused on text data [5][12].

There are some of the efficient approaches to automatically group relevant web videos into events, and then mine the relationships between events. Visual information suffers from semantic gap [7]and user subjectivity problems so that to find NDK is difficult [10]. The textual information can be noisy, spelling mistakes, and sometimes missing character in the text. Therefore, the problems can make the unsatisfactory results in web video event mining. Such challenges motivate us to utilize textual features for web video event mining[4]. To address this need, web video event mining approaches have been developed. a novel framework can be proposed for integrating the visual temporal information and textual distribution information for web video event mining.

After feature selection and tag relevance learning by neighbor voting, MCA is used to extract the text

from each NDK in event similarity [3] with the assistance of textual information. Next, both cooccurrence information and visual near-duplicate feature trajectory induced from NDKs are used to detect the similarity between NDKs and events. Finally, in order to integrate visual and textual information for event mining.

II.EXISTING SYSTEM

Several content-based video search engines had been proposed recently. The system retrieved more than 37.000 lecture videos from different resources such as YouTube, Berkeley Webcast etc. The search indices were created based on the global metadata obtained from the video hosting website and texts extracted from slide videos by using a standard OCR engine. Since they do not apply text detection and text segmentation process, the OCR recognition accuracy of their approach is therefore lower than the systems.

Furthermore, by applying the text detection process they are able to extract the structured text line such as title, subtitle, key-point etc. that enables a more flexible search function. An entity recognition algorithm and an open knowledge base are used to extract entities from the textual metadata. As mentioned before, searching through the recognition results with a degree of confidence, they had to deal with the solidity and the consistency problem.

DEMERITS

- The data show that over 100 h of videos are uploaded to YouTube every minute, and six billion hours of videos are watched each month.
- The users had to browse many web videos in the returned list and even watch most of the videos. This is not only time consuming, but also difficult if thousands of web videos are returned by a search engine.
- A large number of web videos are returned from video search engines mostly according to the text relevance, there exist many noisy items in the search results, which diminishes the mining performance.
- These facts demonstrate the users cannot grasp the major events available from searching video databases.

III. PROPOSED SYSTEM

The proposed framework for web video event mining consists of four stages:

1. Keyframe Conversion and Terms Preprocessing
2. Cooccurrence Detection and Feature Trajectory
3. Text Feature Selection
4. Event Mining

The textual information's are noisy so that preprocessing method is used to filter the noise from text. After getting the correlation between terms and events through visual neighbor information, MCA is used to calculate the similarity between each NDK and event through textual distribution information in NDKs.

For the visual information, the video are composed of several keyframes. First to convert the videos into keyframes, after the Keyframe conversion to find the NDK keyframes within video information and then visual near-duplicate feature trajectory [2] are applied to improve the robustness and accuracy of the visual features in the keyframes. Finally, the content-based visual temporal information and textual distribution information are integrated through the proposed model.

Event mining is achieved within each topic[5]. The input of the framework is the web videos and user query. As a result, the similarity between each NDK and each event is calculated, and every NDK is assigned to the event with the largest similarity. The output is the classified events.

IV.SYSTEM ARCHITECTURE

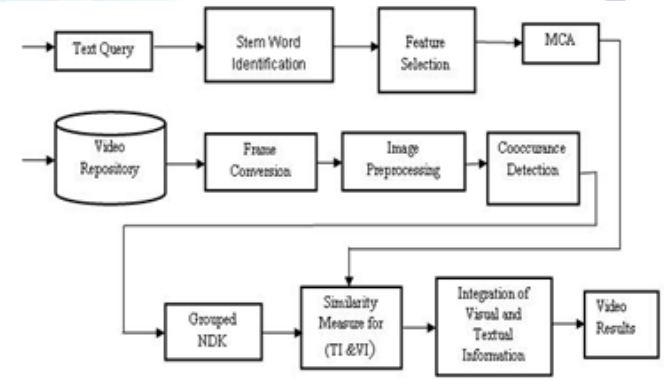

Figure: 1System Architecture

V.SYSTEM DESIGN

A. Keyframe Conversion and Terms Preprocessing
B. Cooccurrence Detection and Feature Trajectory
C. Text Feature Selection
D. Event Mining

A. KEYFRAME CONVERSION AND TERMS PREPROCESSING

In this module, Videos files are retrieved from video repository and then convert the videos into key frames. First find the video length, frame rate it is useful for converting the videos into key frames. After shot boundary detection, the middle frame is extracted in the video as the key frame for the shot. Each video can be represented by a sequence of key frames. NDK is a group of key frames that are visually similar, but different because of the variation during the acquisition time, lens setting, lighting condition, and editing operation.

The input of the framework is text query. The text query may be lengthy so the important word can be identified in the text query.The Word Stemmer is used to identify the stop words in the given text query. Then, the identified stop words are eliminated. Finally the stem words can be obtained based on the word the featured can be matched in the video keyframes.

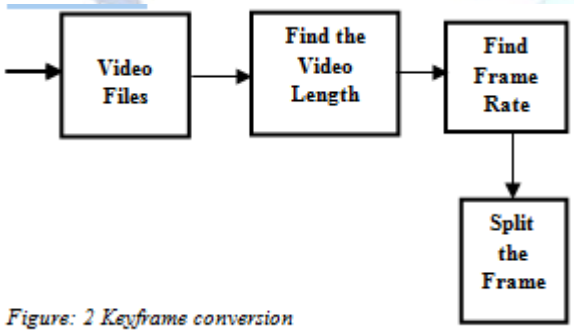

Figure: 2 Keyframe conversion

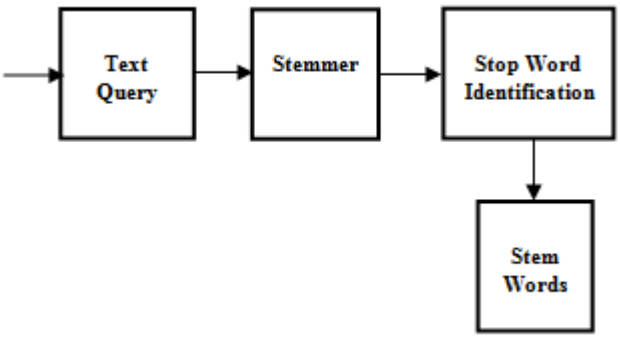

Figure: 3Terms Preprocessing

B.COOCCURRENCE DETECTION AND FEATURE TRAJECTORY

In this module, after the Keyframe conversion and then calculate the similarity between the neighborhood keyframes in the video event.Finally obtain the similarity and then find the mean value between the sequences of keyframes/images. The result of this module is to extract clear keyframes from the video shot. That Keyframe is said to be cooccurrence (Near Duplicate Keyframe). The text can be identified in the cooccurrence keyframes. After that feature trajectory [6] method is applied to improve the robustness of the visual features.

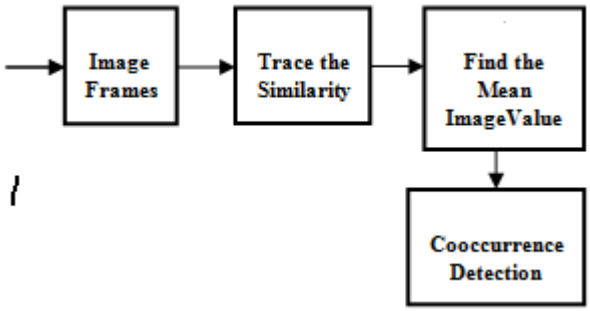

Figure: 4 Cooccurrence Detection

C.TEXT FEATURE SELECTION

In this module, the stem words are obtained and then text can be identified by using lee sign feature algorithm. The text, numbers can be pre-trained and stored in the repository. Based on this features to measure the similarity between the text.The video are retrieved from search results according to the text and tag relevance [11], to locate the bursty period due to video upload time [8].

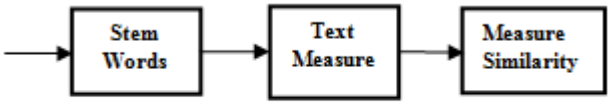

Figure: 5 Text Feature Selection

D.EVENT MINING

In this module, the text can be extracted from the Keyframe. Then user text query can be matched with extracted Keyframe text. If the text can be matched means it retrieve the videos from the video repository by using proposed model. Finally to obtain the video files based on the user text.

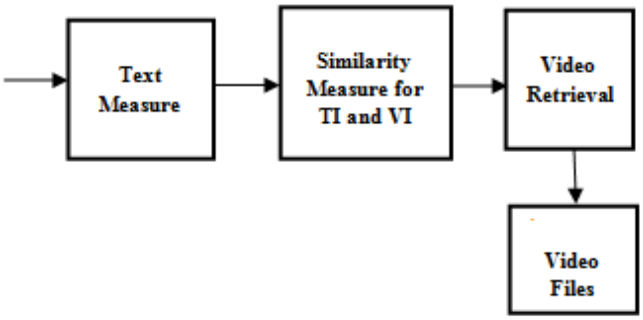

Figure: 6 Event Mining

VI. CONCLUSION AND FUTURE ENHANCEMENT

The characteristics of web videos such as unavoidable video editing error, NDK detection and noisy textual information, so the above characteristics are reducing the web video mining performances, it is a challenging task. In this paper a novel framework can be proposed for integrating the visual temporal information and textual distribution information and also solve the noisy, unavoidable video editing and NDK detection problems. The result of this framework is to retrieve the reletated web video based on user text from the more number of available video without searching every video within the video archive. The future work, is improved the text feature learning mechanism, such as sparse representation and Fourier transform algorithm is used to extract the text from blurred images due to fast motion video and retrieval.

REFERENCES

[1] Statistics (2016).[Online]. Available: http://www.youtube.com

[2] Chengde Zhang, Xiao Wu, Member, IEEE, Mei-Ling Shyu, Senior Member, IEEE, and Qiang Peng, "Integration of Visual Temporal Information and Textual Distribution Information for News WebVideo Event Mining", Feb, 2016.

[3] C. Zhang, X.Wu, M.-L.Shyu, and Q. Peng, "A novel web video event mining framework with the integration of correlation and co-occurrence information," J. Comput.Sci. Technol., vol. 28, no. 5, pp. 788–796, 2013.

[4] X. Wu,C.-W. Ngo, and Q. Li, "Threading and autodocumenting news videos: A promising solution to rapidly browse news topics," IEEE SignalProcess. Mag., vol. 23, no. 2, pp. 59–68, Mar. 2006.

[5] J. Cao, C.-W.Ngo, Y.-D.Zhang, and J.-T. Li, "Tracking web video topics: Discovery, visualization, and monitoring," IEEE Trans. Circuits Syst. VideoTechnol., vol. 21, no. 12, pp. 1835–1846, Dec. 2011.

[6] Q.He,K. Chang, and E.-P. Lim, "Analyzing feature trajectories for event detection," in Proc. 30th ACM Int. Conf.Res.Develop.Inform.Retrieval,2007, pp. 207–214.

[7] C. Xu, Y.-F. Zhang, G. Zhu, Y. Rui, H. Lu, and Q. Huang, "Using web-cast text for semantic event detection in broadcast sports video," IEEE Trans.Multimedia, vol. 10, no. 7, pp. 1342–1355, Nov. 2008.

[8] X. Wu, Y.-J.Lu, Q. Peng, and C.-W.Ngo, "Mining event structures from web videos," IEEE Multimedia, vol. 18, no. 1, pp. 38–51, Jan. 2011.

[9] W.-L. Zhao, X. Wu, and C.-W. Ngo, "On the annotation of web videos by efficient near-duplicate search," IEEE Trans. Multimedia, vol. 12, no. 5,pp. 448–461, Aug. 2010.

[10] J. Yuan, Y.-L.Zhao, H. Luan, M. Wang, and T.-S. Chua, "Memory recall based video search: Finding videos you have seen before based on your memory," ACM Trans. Multimedia Comput., Commun. Appl., vol. 10, no. 2, pp. 1–21, 2014.

[11] X. Li, C. G. Snoek, and M. Worring, "Learning social tag relevance by neighbor voting," IEEE Trans. Multimedia, vol. 11, no. 7, pp. 1310–1322, Nov. 2009.

[12] K.-Y. Chen, L. Luesukprasert, S.-C. T. Chou, "Hot topic extraction based on timeline analysis and multidimensional sentence modeling," *IEEE Trans.Knowl. Data Eng.*, vol. 19, no. 8, pp. 1016–1025, Aug. 2007.