

A Survey Paper on News Web Video Event Mining

Aggie Varghese ^[1], Smitha M. Jasmine ^[2]

Department of Computer Science and Engineering
India

ABSTRACT

Due to the explosive growth of web videos, it becomes great challenge of how to efficiently browse lakhs or even millions of videos at a glance. Given a user query, social media web sites usually return a large number of videos that are diverse and not relevant. Exploring such results will be time-consuming and thus degrades user experience. Based on these observations, an event mining solution should provide the users a quick overview for each topic with a reduced browsing time. There are different methods for news web video event mining and are described in this paper.

Keywords :- Cooccurrence, multiple correspondence analysis (MCA), near-duplicate keyframes (NDK), news web video event mining, visual feature trajectory.

I. INTRODUCTION

News web videos are mainly composed of visual and textual information. Visual information contains semantic gap and user subjectivity problems, and therefore, using either visual or textual information alone for news web video event mining may lead to unsatisfactory results. In order to overcome these shortcomings, both visual and textual features are utilized for web news video event mining. For visual information, some important shots are frequently inserted into videos as a support of viewpoints, which carry useful information. Since there is unique role of near-duplicate keyframes (NDK) in the news search, topic detection and tracking (TDT) and copyright infringement detection, these duplicate keyframes/shots are clustered to form different groups according to visual content. Such groups are similar to the hot terms in the text field. Here, each cluster is called an NDK group, which can be used to group videos with similar content to the same events.

II. RELATED WORKS

The various methods used for news web video event mining are described here.

Q. He, K. Chang, and E.-P. Lim, in [1] “Analyzing feature trajectories for event detection,” proposed a method based on the feature trajectory using text words. In this paper, they 1) spectral analysis is applied first to categorize features for different event characteristics: periodic and aperiodic, important or less-reported. 2) modeled periodic features with Gaussian mixture densities and aperiodic features with Gaussian density and, and detected each feature’s burst by the truncated Gaussian approach; 3) proposed an unsupervised greedy event detection algorithm to detect both aperiodic and periodic events. The highly sets to the orrelated word features are grouped to events by mapping the word video sets. The performance of FT_T is poor. This c

is a challenging mission to mine the events under the noisy and diverse social web scenario.

J. Yao, B. Cui, Y. Huang, and Y. Zhou, in [2] “Bursty event detection from collaborative tags,” proposed a technique based on visual Near Duplicate Keyframe -level clustering. In this paper, they proposed a new method to detect bursty tagging event, which captures the relations among a group of correlated tags. These tags are either bursty or associated with bursty tag co-occurrence. This kind of bursty tagging event generally corresponds to a real life event. The events are profiled with more comprehensible and representative clues. The proposed method is divided into three stages. As the first step, they exploit the sliding time intervals to extract bursty features, and graph clustering techniques is adopted to group bursty features into meaningful bursty events. It helps to detect the relationship among tags with bursty tag cooccurrence. In this method, the number of NDKs among videos is limited. Each event is composed of different kinds of visual scenes, while CC_V just groups one scene.

X. Wu, Y.-J. Lu, Q. Peng, and C.-W. Ngo, in [3] “Mining event structures from web videos,” proposed T+V which is a fusion method. This paper explores the issues of mining event structures from Web video search results using burst detection, text analysis, clustering, and other techniques. It first applies feature trajectory to the visual field. Then it tries to mine events on the basis of text cooccurrence and feature trajectory. In this method the visual near-duplicate feature trajectories of NDKs are not consistent. It misses those low-frequency terms and NDKs which are common for a large number of videos.

C. Zhang, X. Wu, M.-L. Shyu, and Q. Peng, in [4] “A novel web video event mining framework with the integration of correlation and co-occurrence information,” proposed a method inorder to mine the correlation between

events and NDKs by using the distribution characteristics of the terms. In this paper, in order to improve the performance of web video event mining they propose a novel four-stage framework. The first stage is data preprocessing. Multiple Correspondence Analysis (MCA) is then applied to explore the correlation between terms and classes. It targets to bridge the gap between NDKs and high-level semantic concepts. Next, the similarity between NDKs and classes are detected using co-occurrence information. Finally, through negative NDK pruning and positive NDK enhancement both of them are integrated for web video event mining. It can bridge the gap between NDKs and terms. In this method, multiple languages, synonyms, and the number of videos in each NDK are problems.

Chengde Zhang, Xiao Wu, Mei-Ling Shyu, and Qiang Peng in [5] “Integration of visual temporal information and textual distribution information for news web video event mining”, proposed a method which uses both the neighbor stabilization process and MCA similarity measure to generate the textual distribution information. It can better explore the degree of correlation between different terms and events. The visual near-duplicate feature trajectory, i.e., the time distribution information of an NDK, is integrated with the NDK-within-video information (cooccurrence) as the visual temporal information to cluster more NDKs belonging to the same event.

III. COMPARISON BETWEEN RELATED WORKS

To evaluate the performance of the web event mining, we use the Precision (P), Recall (R), and F1 measure (F1), which are defined in (1)–(3), respectively

$$\text{Precision} = |B_i^+| / |A_i| \tag{1}$$

$$\text{Recall} = |B_i^+| / |B_i| \tag{2}$$

$$F1 = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{3}$$

where B_i^+ is the number of correctly grouped positive videos for cluster A_i , and B_i is the number of positive samples in the ground truth. Since F1 considers both Precision and Recall values, it is mainly used to evaluate the performance.

Methods	Precision	Recall	F1
Analyzing feature trajectories for event detection [FT_T]	low	low	low
Bursty event detection from collaborative tags [CC_V]	high	low	low

Mining event structures from web videos [T+V]	Better than FT_T	Better than FT_T and CC_V	Better than FT_T and CC_V
A novel web video event mining framework with the integration of correlation and co-occurrence information [MCA]	low	Sometimes high	Can be low or high compared to FT_T, CC_V and MCA
Integration of visual temporal information and textual distribution information for news web video event mining	Better than FT_T, T+V and MCA	Most of the time high	high

IV. CONCLUSION AND FUTURE WORKS

This paper compares different methods used for news web video event mining. The method used in integration of visual temporal information and textual distribution information for news web video event mining can accurately identify more relevant news web videos. Since CC_V method focuses on visual content, it might miss many news web videos without the duplicate visual content. While applying the method of FT_T to the visual NDK level, different scenes of the news web videos can be grouped together which indicates that the visual near duplicate feature trajectory can deal with news web video event mining with diverse visual scenes. However, the NDK detection error problem may lead to inaccurate visual near-duplicate feature trajectories. Therefore, the instability characteristics would mine unrelated news web videos pertaining to other events. MCA focuses on the textual information. For some news web videos, in order to attract more attention, the users might add some words in the title/tag. Hence, some unrelated keyframes are falsely considered as relevant information. The visual information are used for grouping the relevant videos into events. The texts/terms are relatively general, broad, and noisy. Hence, the integration of both visual and textual information can lead to good results. Finally, clustering with the textual and low frequent information could yield bring more relevant videos. The text information extracted from the keyframes of the videos can also be taken as the textual information for the event mining process.

ACKNOWLEDGMENT

We thank members of Computer Science and Engineering department of College of Engineering and Management, Punnapra, for their valuable support and feedbacks.

REFERENCES

- [1] 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.
- [2] J. Yao, B. Cui, Y. Huang, and Y. Zhou, “Bursty event detection from collaborative tags,” *World Wide Web*, vol. 15, no. 2, pp. 171–195, 2012.
- [3] 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.
- [4] 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.
- [5] Chengde Zhang, Xiao Wu, Mei-Ling Shyu, and Qiang Peng in [5] “Integration of visual temporal information and textual distribution information for news web video event mining”
- [6] Statistics. (2017). [Online]. Available: <http://www.youtube.com/yt/press/statistics.html>