International Journal of Trend in Research and Development, Volume 4(2), ISSN: 2394-9333
www.ijtrd.com

Semantic Based Video Hierarchy for Browsing Large Video Repositories

[1]M. Deepika, [2]K. Dhanalaxmi, [3]B. Karthika, [4]I. Persi Pamela
[1,2,3]Final Year B.Tech (IT) Student , [4]Assistant Professor,
[1,2,3,4]IT Department, Kingston Engineering College, Vellore, Tamil Nadu, India

Abstract- The number of videos on the Internet is growing explosively, which poses a great challenge on automatic video search. Existing video search engines, such as YouTube, visualize the search results in a ranked format. The simple list structure may be suitable for locating a movie trailer or a music video, but cannot handle queries with complex topic structures so it creates difficulties for the users to locate their needed information. In this case, a semantically structured hierarchy with nodes covering all the critical topic facets becomes a much better choice. Obviously, manually creating topic hierarchies for every query is impractical simply due to the great number of the search topics. We use a Prototype-based Hierarchical Clustering algorithm to create an automatically generate hierarchical structure by processing the video descriptions along with their semantic information. We extract this information from a semantic database using a sub-ontology concept because we can't have the whole semantic database in our website.

Keywords: *Semantic Database, Sub-Ontology, Prototype- Based Hierarchical Clustering*

I. INTRODUCTION

Now-a-days videos on internet is evolving. Everybody uses the video search engine for getting information. Most of the video search are done by matching a query keyword with the video description given by the uploaders. But some of the uploaders to get the viewers they will give wrong description for the video. So viewers will not get their relevant information. The search engine creates a keyword index structure for enhancing the search performance of the video. An indexing system without controlling the vocabulary is the Natural language indexing. Keyword is a catchword in the description given by the uploaders. Creating a keyword index for searching video is not sufficient complex query structure. For instance, if we search for the keyword MJ_rehersal, the search results will have multiple videos related to MJ such as his concerts, performance, etc, there will also have the results of MJ_rehearsal but it will be difficult for the users to access the correct video. The videos are displayed by using the keyword indexing.

Due to significant difficulty of creating a high quality video search we use prototype based hierarchical clustering algorithm for processing with their semantic information as a basic for visualize the search result with a high quality. Semantic database is a huge database which contains large size of data. It is difficult to extract the whole semantic information. So we use sub-ontology technique for extracting information from the semantic database.

Sub-Ontology is extracting the information from the existing database. By using the sub-ontology technique the data is extracted from the large database and the relationship is made between the video and the name of the video.

II. RELATED WORK

The video search engines uses a feedback based re-ranking scheme for updating the keyword index as a process of understanding the querying users. Initially they used the keyword index based on the description given by the user at first, which is not sufficient. And then they used feedback and number of views of the users and they matched it with the query keyword to find the relationship of the video and the query keyword. Based upon the relationship of the keyword and video the index structure is updated and re-ranked. For instance, if a user searches a video and views it, then interaction for the video and keyword is done. So the ranking for the video will increase.

Yu-Gang Jiang Wang, Qiang Wang, Wei Liu and Chong-Wah Ngo[1] In this they have introduced hierarchical visualization approach for the search results, so that the viewer can quickly understand the multiple results of a query topic in a well organized manner. In this the textual descriptions are arranged in hierarchical manner. Here useful topics are not covered by the hierarchy automatically to ensure the good result. Santhana Krishnamachari, Mohamed Abdel-Mottaleb [2] Digital video clips are more popular nowadays due to the more availability of user devices that captures more videos and uploading it in social websites. So the digital contents have been increasing explosively, so there is a need for the user friendly tools. They are using some hierarchical clustering algorithm to navigating the tree structured result.

Pavel Calado, Marco Cristo, Edleno Moura, Nivio Ziviani Berthier Ribeiro-Neto Marcos Andr´e Gonc¸alves [3]- In this the link information can be used to improve the results of the classification for a web collections. They used Bayesian network model to combine the link- based and content- based information. But the links which are internal to the collection do not provide sufficient information for classifying a document.

Zhao-Yan-Ming, Kai Wang, Tat-Seng Chua[4]- They used the prototype based hierarchical clustering algorithm for organizing the web collections and it simultaneously solves the categorizing and interpreting the clustering result problem for navigation. There is a lack of efficiency in the proposed PHC algorithm and applicability in multimedia collections. Jiajun Wang, Yu-Gang Jiang, Qiang Wang, Kuiyuan Yang, Chong-Wah Ngo [5]- This paper proposes a technique to improve the efficiency of browsing complex query topics by organizing the search results into semantical structure hierarchy. But there is a need to remove the irrelevant information and also add the useful information.

Chien-Chung Huang, Shui-Lung Chuang, Lee-Feng Chien[6] - the human interaction is very important for the hierarchical

classification. So they used a technique called LiveClassifier that can automatically train classifiers based on user-defined topic hierarchies through web corpora. Xiao Wu, Alexander G. Hauptmann, Chong-Wah Ngo[7]- Many web search results are dependent on video tags and text keywords which results in displaying duplicates and near duplicates of the search results. So, to overcome they using the hierarchical clustering, but it is challenge to detect the near duplicates in upcoming web 2.0.

Song tan, Yu-gang jiang, Chong-wah ngo[8]- The videos are arranged in an very structured way. The videos are placed accurately by using the semantic hierarchy. For searching the suitable videos they use the three criteria: relevance, uniqueness and diversity. But not all relevant videos are placed in a hierarchy, and complex queries cannot be matched with the Wikipedia pages. Alex Hindle, Jie Shao, Dan Lin, Jiaheng Lu, Rui Zhang[9]- It clusters the videos by using the various evidences from the variety of information sources such as tags, description, titles etc. But only partial video results are relevant to the query. Some are mixed with the irrelevant query results.

III. SEMANTIC VIDEO HIERARCHY

Instead of creating a index structure for searching we use a hierarchy structure for enhancing the video results. The hierarchy is constructed by extracting the key points from the description and find the hierarchy relationship between the key points. We also use a semantic database for finding the relationship between the keyword and we finally build a classification process for grouping the documents. The group of classified documents is place under the hierarchical tree and when a user search a document the group is given as a search result. We also visualize the video document in hierarchy structure so that user can get videos which they have searched for.

User can visualize document in a hierarchical structure so it reduce the need of searching the entire document. It also reduces the irritation of the user while compared with the existing index structure system. Thereby, user can perform search and viewing functionalities so that they can find video what they want.

A. System Architecture

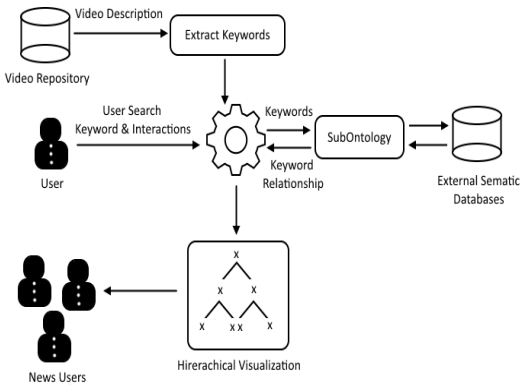

Figure (a): System architecture

The figure (a) shows the architecture of the semantic based video hierarchy browsing the large video repository. In the video repository the multiple videos are stored. The video repository is user based module. The user searches the video by using the keyword. The system extracts the keyword from the video description given by the user. The keyword from the description and the user searched keyword are processed. The keywords are processed further by using the sub Ontology technique and extract the video from the external semantic databases. Only related videos are extracted from the semantic databases by using the sub ontology technique. The user is interacted with the keyword and the relationship is made between the keyword and interaction. The keyword is mapped to the video. By using the keyword relationship the videos are arranged in an hierarchical format. The user can view the videos in an hierarchical visualization.

If new users want to search a video, the video which has the high ranking i.e the video which has more interaction with the keyword is placed first. Since the videos has hierarchical visualization, it will be easy for new users to view the video which they have searched for.

B. Module Description

1. Extracting Keyword:

Each video are associated with its related textual description given by the uploaders. The system extracts the keywords from the given description. The system also maintains the database with the predefined keyword dictionary. The keywords are extracted from the textual description and the video is searched by using the keyword dictionary database. The frequency of the keyword is calculated which is given by the textual description. The frequency is calculated by number of views by the users. The views and keyword matching is considered for the frequency calculation.

2. Construct Index

Based on the extracted keyword and their frequency we construct an weighted index structure. A threshold value for keyword frequency is set. Threshold value is the critical value for the video and the keyword.

If the keyword is least related to the query search then the video has low threshold value. And if the video has high frequency i.e. more number of keyword interaction then its threshold value is high. The value below the threshold is removed from the index structure. Index structure is updated when video description is updated and by the user search interaction which is done by next module.

3. Monitoring User Interaction

This module is used to extract relationship between a new keyword and video during the user search and selection. User select a video from the search result generated for their keyword search. From this we assume that the selected video is more associated to their search keyword. Based on that we increase the weightage of the keyword, if keyword is not found we add that keyword to the index.

4. Extracting Relationship

Existing system extract prototype hierarchies from a Wikipedia pages but we are constructing of hierarchy based on semantic structures. Semantic structures contain keyword and their relationship sketched in a tree structure, these relationships are mostly made by human. They are huge in size so we extract a sub tree from the big semantic tree based on our constructed index structure

5. Construct Hierarchy

After extraction of the keyword and their relationships we sort our keyword into different level according to their relationship. These sorted linear keyword set are merged to form a hierarchy tree where each sub node are our extracted keywords. Based on the keyword weights in constructed index videos are inserted into the respective node in the hierarchy.

CONCLUSION

We have introduced the semantic based video browsing in the large video repository. The videos are uploaded and it automatically ensures the search results. The search results are automatically updates and the related videos are displayed. The videos are arranged in a hierarchical manner to avoid the user irritation while searching. In this the videos are automatically removed if the videos have fewer thresholds and it is arranged in the index structure. After the interaction of the videos the related videos are arranged in a hierarchical manner. The videos are extracted from the semantic database and also by using the subontology technique. The videos which are extracted are arranged in a hierarchical format.

References

[1] Yu-Gang Jiang, Jiajun Wang, Qiang Wang, Wei Liu, and Chong-Wah Ngo, "Hierarchical Visualization of Video Search Results for Topic-Based Browsing" *in proc of IEEE Multimedia*, vol. 18, no. 11, november 2016

[2] Santhana Krishnamachari, Mohamed Abdel-Mottaleb, "Image Browsing using Hierarchical Clustering" *in Proc. IEEE Int. Symp. Comput. Commun.*, 1999.

[3] Pavel Calado, Marco Cristo, Edleno Moura, Nivio Ziviani Berthier Ribeiro-Neto Marcos Andr´e Gonc¸alves, "Combining Link-Based and Content-Based Methods for Web Document Classification" *in Proc. ACM 12th Int. Conf. Inf. Knowl. Manage.*, 2003.

[4] Zhao-Yan Ming, Kai Wang and Tat-Seng Chua,"Prototype Hierarchy Based Clustering for the Categorization and Navigation of Web Collections" *in Proc.Prototype Hierarchy Based Clustering Categorization Navigation Web Collections*, 2010.

[5] Jiajun Wang, Yu-Gang Jiang, Qiang Wang, Kuiyuan Yang, Chong-Wah Ngo, "Organizing Video Search Results to Adapted Semantic Hierarchies for Topic-based Browsing" *in Proc. ACM Int. Conf. Multimedia*, 2014.

[6] Chien-Chung Huang, Shui-Lung Chuang, Lee-Feng Chien, "LiveClassifer: Creating Hierarchical Text Classifers through Web Corpora" *in Proc. 13th Int. Conf. World Wide Web*, 2004.

[7] Xiao Wu, Alexander G. Hauptmann, Chong-Wah Ngo," Practical Elimination of Near-Duplicates from Web Video Search" *in Proc. ACM Int. Conf. Multimedia*, 2007.

[8] Song tan, Yu-gang jiang, Chong-wah ngo,"Placing Videos on a Semantic Hierarchy for Search Result Navigation" *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 10, 2014.

[9] Alex Hindle, Jie Shao, Dan Lin, Jiaheng Lu, Rui Zhang," Clustering Web Video Search Results based on Integration of Multiple Features" *in WWW J.*, vol. 14, no. 1, 2011.