Deepika Sagar

Department of Computer Science and Engineering Manav Rachna International Institute of Research & Studies

Jwngfu Brahma

Department of Computer Science and Engineering Manav Rachna International Institute of Research & Studies

Shobha Tyagi

Department of Computer Science and Engineering, Manav Rachna International Institute of Research & Studies

Abstract

Information retrieval is a field that relates to the way of collecting, managing and retrieving data using different techniques. In this paper we have discussed the background activities through which the information is provided to the user in the form of multimedia content. Effective retrieval of information includes use of detectors and trackers that tracks the information on the basis of the logs of the previous searches. A video search engine working depends upon the operations such as crawling, indexing and finally ranking. Various ranking based criteria are involved to retrieve the best result to the user from the large quest of data. The main purpose of the use of various techniques within a search engine is to support and manage a large number of audio-visual contents. Search engine optimization is a well practice technique to provide the ranking and assigning the appropriate position to the sites over a search engine. This paper represents the various works done by different researchers in the field of information retrieval and the future use of the video search engines along with the current trends and technologies used by different video search engines.

Keywords:

Video search engines, Detectors, Trackers, Word recognition, Ranking, Search Engine Optimization

Introduction

With the growing advancements in the field of science and technology, the digital networks and the related sectors are enhancing with a fast speed. This give rise to the need of a proper technique for the retrieval of digital information from a quest of data. Being in the twenty-first century, we have an option of choosing from several tools to suit our basic day to day needs, Google Search, Yahoo Search, MS Bing and DuckDuckGo being the most widely used ones. With the help of these smart search engines one can find any information publicly available online, around the world. Over the years the Multimedia Broadcasting Platforms have gone under huge transformation. What earlier was limited to the Studios and TV Channels is now within the reach of everyone, anybody can create an account in any of the online broadcasting services and literally start uploading the contents right away. Retrieving the desired data from such a huge volume of visual data without proper visual identification mechanism is a tough job, as every video search engine available online depends only upon the textual

description of the visual data. Even the smartest available search engines depend only upon the speech recognition which obviously isn't always accurate, leaving out the visual aspect that comprises around 60% of the real meaning of the data. These problems can be only solved when we teach the machines to sort the visuals like we do. This can be made possible with the help of proper sentence tracking, word detection or object detection technique that is applied for tracking the activities in a specific video clip. The video search engines take up the query in the form of a sentence and a sentence consist of different phrases therefore different detectors and trackers are used to detect if noun occurs or a verb occurs or adverb etc. In many cases two different sentences can have the same set of words but the phrased sentences may have different meaning that's why there is a need to have a system that efficiently detects the difference between the two. The problem can be well explained with the help of an example: "The thief threatens the lady" and "The lady threatens the thief", though these are two different sentences with different meaning but contains

the same set of words and if one doesn't have proper detecting technology it will end up showing same result for these sentences. This is the reason for understanding the important to develop a system that not only detects and understands the difference between the two or any number of sentences but also come up with a meaningful output. This can effectively be done with the help of proper use of technology that includes use of Compositional semantics [1] to construct the meaning of a sentence from the meaning of its words to detect the difference. Similarly, one uses the Concept detection [2] for detecting high level features (also called concept) in videos. Robust video retrieval technique[3] Includes the algorithm that attempts to match the description of the information as provided by user from a collection of documents already present in their database.

ELEMENTS BEHIND SEARCH ENGINES

The objective of this section is to discuss the elements that serves as basicity for constructing as robust, well managed and easy to handle video search engines. The main necessity in a video search engine is to understand the query that user provide over the search bar. The elements provided below helps in such process.

Detectors

An object detection technique [4] is a computer technology associated with computer vision and computer image processing. Detectors deals with the recognition of the occurrences of semantic objects in a digital image and media. It is also utilized in fields like face recognition and pedestrian detection and many others. Its vast utility and utilization make this technology important for security related fields and sectors. Object detection technology applications includes computer vision, including picture recovery and image retrieval, video surveillance, tracking various objects like tracking a car on a street, tracking movements of objects like a ball in a video. The basic concept includes the recognition of objects.

Tracking

To search for the recordings (media) which portray a sentence [5], a search engine first track all the object that take part in that particular event depicted by that sentence. Sentence detection technique is used to recognize each frame of the video. An object detector runs on each frame and detects all the objects within each frame. Ranking is done on each track that can be used in future for better retrieval of information. A log of previously asked tracks and currently used tracks were maintained by the servers in order to keep track of various activities occurring and also for future use and requirements. The future requirements includes the use of all the tracked information to maintain a good quality of service and in increasing the efficiency of the search engine. We generally require tracker for two reasons

first there exist is a bunch of object detection technique in the market so to get the most optimal solution and second not all the object detection technique is reliable.

Word and Sentence Recognition

It refers to the capability of a search engine to perceive composed words accurately and effortlessly. Word recognition technique [6] is a way of recognizing each words that a group of words referring to. A logical set of letters gives a logical word and a logical set of words together forms a logical sentence. Sentence recognition technique refers to the technique of recognizes each word of a sentence. Algorithm is capable of finding out the number of times verb occurs within a sentence and the number of times the noun occurs the adjective and other necessary information to provide the user with the most optimal result. A sentence consists of words and each word defines a set of objects related to it that is probably worked upon by the search engine. A search engine should be capable of differentiating the difference between the sentences containing the same set of word but depict different meaning like one happening in the following sentences "THE THIEF THREATENS THE LADY" and" THE LADY THREATENS THE THIEF".

Retrieval

The final step is to retrieve the relevant results out of the search index that can be forwarded to the user asking for it. The final output shown to the user is the list of relevant searches sorted in order of their degree of relevance and ambiguity. The output is retrieved from the index that the search engine has created using various tracking techniques. Logged data also helps a search engine in providing the quality of information to the user. That is why a large sum of money is spent by the sited in order to maintain and secure these logs. It is necessary to provide the optimal output to the user in order to maintain the trust of the users. Thus the final output is approximately the one that the user is expecting.

WORKING

The main purpose of such search engines is to support a large number of audio-visual content and adoption of various techniques to manage it. Due to digitalization and advancement in technology there is a need to support this type of technology for various purposes like study, research and many more. According to researches people grab or understand things bitterly with the help of an image or from a video as compared to text only information. This is the reason that the government over the world are investing more on making books, papers, navels digital so that more and more people can assess it anytime and anywhere and hence it reduces the pain of carrying heavy book from one to another. The working of search engines depend on the operations that includes crawling, indexing and ranking.

First we discuss Crawling in detail. (i) Crawling: The search engines consist of web crawlers[7] or web spiders that are basically the software robots that browse billions of pages found on the net and maintain a track record of the valid searches. Spider basically analyzes the metadata and the keywordsof the page and that provided by the user and keeps on building the list of content. (ii) Indexing: After the crawling [7]process, all the relevant results and related outputs are coupled up by the crawler. Search engine decode, sort and store each result file [8] in a different location generally on a virtual warehouse, that are sorted according to the phrases and the keywords defining them for the easy retrieval of the information. Spider reports back the result to the search engine where they are shown to the users in action. A search engine uses different algorithms [9] to search and retrieve the relevant information from the index. Web spiders not only crawls the relevant information but also analyse each result and rank them. (iii) Ranking: Once the information is set in the database it is indexed on the basis of the factors including number of times page visited, duration of each visit, number of backlinks and many more to determine how relevant the information is [10]. It is to be noticed that the ranking is done to know the usefulness of both the search engine and its searching capabilities. There are billions of videos over the internet and out of that thousands of videos are there that contains same words or same phrase, thus making it even trickier to get the relevant data out of such an information pool. Ranking make the output simple by indexing the information on the basis of their relevance. The relevance of a video depends on various factors on the basis of which optimum solution is obtained and is presented to the users. Further the indexing can be based on different parameters. Next we discuss it in details.

Indexing by views counts

Sorting the results based on the total number of the views that shows the popularity of the video. Considering the famous site YouTube whose algorithm calculates the total number of views only if that particular video is watched for atleast 30 second. YouTube also check for any misleads that probably includes buying the views by the people. YouTube here uses a technique where it freezes the video at a particular views say 300 views, at this point its algorithm whether the views are real or not.

Indexing by relevance

Sorting the results based on the integrity and ambiguity of the information. It is the responsibility of the uploader to upload the video the help or entertain people in one or the other way. If the video uploaded doesn't make any sense then it is of no use and various search engines adopt different technique to keep such irrelevant videos at a distance. Viewers comment also helps in considering a video relevance. The relevancy of the information depends on the accuracy, uniqueness and trustworthiness of the data.

Indexing by number of spams obtained

Sorting results based on the number of negative of spams reported against a video. If the viewer find cheated in any way then he has whole right to report a spam against it so that other can be saved. Various search engines adopt different policies against spams in order to believe the users how safe there engine can be. Increased numbers of spams may damage the reputation of a site and under poor circumstances even the sites has to be called off for protecting the users from any unauthorized and damaging action.

Indexing by upload date

Sorting result based on the time of their upload. Results are also sorted according to the seniority in the repository. The criteria include Last hour, Today, This week, This month, This year. Old and less viewed videos are kept at least priority and is sometimes out of the list if not asked at all for a couple of years. Through the uploaded date one can find the popularity of a video by the number of views it gets in lesser time.

Indexing by the duration

Sorting result based on the duration of the videos. The video can be SHORT that includes all the videos less than 6 minutes and LONG including all the videos more than 20 minutes. Viewer can apply filters depending upon the span the wanted their video to be. Short videos are preferred more for business, entertainment and frequently asked point of view and such videos are efficient to express more within less.

Indexing by rating

Sorting result based totally on the ratings that a video achieved through various users. Most rated video is displayed at the very first page that displays its usefulness. Such videos are recommended by the search engines itself. Least rated videos are kept at last or out of the index if not approached by anyone by a couple of years. Higher the rating of a video means more the effective and popular the video information. Poor rating means either the video is frequently viewed or irrelevant from the user's point of view. **Indexing by the number of subscriber**

Sorting the results based on the number of the users subscribing for a specific video. This criterion is totally based on the popularity of a video among the users. The uploader is paid accordingly by the organization based on the number of subscribers, likes and the positive comment that a video got. Higher the number of subscribers means higher numbers of people are interested in that particular video. Less subscribers means either the video is in its initial state or is less required by the people.

Indexing by result type

Sorting the results based on the type of clip that the users are looking for. This includes the categories such as Video that includes short, medium or long video clips and Movie that includes the media of more than 1 hour. The result provided needs to be accurate and useful by the user's point of view to maintain the user's interest and trust. Trust comes with the reliability of the information as well as the usefulness of the information. Result provided in the end to the users needed to be accurate and up to the user's expectancy.

Indexing by features

Sorting the results based on the additional features that user wants to be in their videos. This includes SD that is 'Simple Definition' it contains all the media of average quality, HD that is 'HIGH DEFINITION' it includes all the high quality media, CC or 'CLOSED CAPTIONS' that contains the subtitles of the media, 3D media for getting a 3d experience, LIVE that includes the media happening live and PURCHASED it include all the media is only made available to the users if they purchase from a particular trusted sites.

Search ENGINE OPTIMIZATION (SEO)

Search engine optimization is a technique that is used to send signals to the search engine that the specific webpage is worth showing. It is basically a process of proving to search engine that you are specific site that is sending the signals is the best site, most trusted, most authoritative and unique that a search engine can give to their researchers. It contains different criteria on which web content are tested among which the three most important criteria include the quality, the trust and the authority criteria.

Quality

It includes how good the content of a website is. The main purpose of having a good quality is to connect to as many people as possible. The quality can be tracked by the records of the previous tracks and modified as per the need of the system to increase the site popularity. This popularity will help in making the site easily accessible to the mass by giving it the specified ranking and place within a search engine. Quality represents the factor that helps in distinguishing the same set of information present on thousands of sites. It is important to have a unique set of information on every topic that a site contains but such situations are sometimes found difficult to create that make it important to represent the data in a well-established manner that can be easily understand by the users if compared with other sites. It's the features like uniqueness and usefulness that not only helps in building the quality but also the performance of a site that makes it different from

Pacific Business Review International

others.

Trust

It depicts how trustworthy a website is. It is checked on the basis of the reviews. Poor review means less trustworthy. Trustworthiness can be increased by connecting to the highly authoritative websites industries link. Other means includes through articles, blogs and many more. Trust is also established through the relevancy of the information provided to the user. The information that a site provides needs to be relevant, up-to- date and the most important it has to be the quality information that must justify the query of the user provided to the search engine. This means that the meaning of the user query is well understood and implemented to find the appropriate result. There is no place for irrelevant or unauthenticated data or facts on a secure platform that helps in building the trust of the people.

Authority

It is a step to show that the website is the most popular website. This is done first by creating fan base that can be effectively built through social media, blogs, getting other website linked by getting good comments. All this will give the site administrative controls and powers that help in building the base of the site stronger. Stronger the base leads to stronger framework on which the site is built. Authorities help in forming an authorized platform to the users where the entry of any unauthorized person is prohibited. Thus it make the website more secure for the people and even more trustworthy where people feel safe to provide their credentials if needed. Such site uses efficient techniques to protect the user's credentials at any means to maintain their reputation.If all the above three criteria are fulfilled efficiently means more efficient one's webpage is and so is the content in that page that at last leads to higher ranking of the page.

Literature Work

Hu et al. [11] worked on semantic video search that mainly focuses on detecting nouns, verbs etc. within a sentence. Authors also used language to search already existing video notation. Snoek et al. [12] worked on interactive search, concept detection and automatic search. The beginning stage for the concept discovery approach is the topperforming pack of-words arrangement which utilizes different colour descriptors, portion- based regulated learning and codebooks with delicate task. Sivic and Zisserman [13] presented their work on retrieving clips and frames both using object detection and query-by-example. He showed a statistical local- feature approach to query-byexample in which a bounding box is placed around an object or a target, and frames in which that object occurs are retrieved. Michael S. Lew and Nicu Sebe[14] Presented Content-based interactive multimedia information retrieval

gives new ideal models and strategies to looking through the bunch assortment of media everywhere throughout the world. This study audits 100+ articles on interactive multimedia information recovery and talks about their part in current research directions that incorporate user studies, high performance indexing, new features and media types and assessment procedures. Sadeghi and Farhadi [15] tracked objects in images using an object detector. They developed a system for a number of multiple interacting objects where before only a single object class is used to consider. This helped them to detect more complex scenarios, such as a person riding a car. Yu et al. [16] provided a unique approach in which he instead of detecting and tracking a single object, a soccer ball, and recognizes all the actions being performed on that object during a soccer match. He examines the position as well as the velocity of the soccer ball and extracted a gross motion feature of that ball this helped him in recognizing events as well as provide him a wide area of research under a small number of domain specific actions that are limited to that single object only. Christel et al. [17], Worring et al. [18], Snoek et al. [19], and Tapaswi et al. [20] presented combination of various technique in one that include text search, noun or verb retrieval that helped them in finding videos with better results. Lin et al. [21] presented an approach to video retrieval with multi-word sentential queries. Kiros et al. [22] worked on an approach that produces text descriptions of still images as well as retrieves still images from a dataset that match multi-word text queries. Shih-Fu Chang; W. Chen; H.J. Meng; H. Sundaram; Di Zhong depicted the rate with which computerized data, especially video, is being created has required the improvement of devices for effective pursuit of these media. Content-based visual inquiries have been fundamentally centered on still picture recovery. They created calculations for computerized video protest division and following, and utilize on-going video altering procedures while reacting to client inquiries.

CURRENT TRENDS AND PRACTICES

The usefulness and efficiency of a search engine depends upon the relevance of its outputs as well as its authenticity. Video search engines are developing at a fast rate and have a bright future ahead due to the growing need. According to Google survey the total world population in 2016 was around 7.3 billion out of which 47% are the internet users. To feed the need of such a large population internet is changing as well as expanding to provide the best to these people. Today video search engines are used for various purposes that includes

Education

Visualizing thing is always better and is more effective than normal reading out something. People across the globe have understood this and is therefore working ahead in the field of making more and more educational websites to teach mass. AR Coden and SW Mak [23] worked on different search engines for the multimedia content. In his paper he talked about the query that comprise of sub queries, each refer to different multimedia content and are used for the search of a collection of multimedia content from the database.

Business

Internet has revolutionized the way of doing business across the globe. Today business either large or small are connecting to the internet to reach to the masses across the globe. Different organizations are showing their work to the world through online seminars, meeting and conferences. JC Monberg [24] worked on the formation of a system that is sufficient enough to provide the business related information. The search engines provide a variety of search option such as search by business name, search by business level, search by business category and many others. Establishment of such search engines provides a platform for the business and related information only.

Entertainment

Video search engines are the means of entertainment for the people as the internet contains 100 millions of videos online on different topics. Hence one can find the videos on the basis of what they want to see in theirfree time. MW Dunn [25] in his paper on interactive entertainment network system talked about a system that offers the user a video-on-demand application. It takes the user choice by providing the user with various options on the basis of which it group together the video content like video games, movies, songs and many more to maintaining a manageable platform for entertainment purposes.

Promotion

In today date more and more people are using internet as a source to promote their services to the people. Their do this through the ads in-between the videos or through a short video clips played either in the beginning or in the end of the video clips. The output of such promotions is more impactful as compared to other means. RM Krapf [26] in his paper on system having videos promotion module tells about having a system that track the user viewing preference within the search engine. The promotional module coupled with that search engine will give the user to select atleast one promotional content to be displayed in order to work further on that search engine.

Conclusion

In the above data we have depicted the ways to deal with video search that takes in data and then processes and analyses it in order to provide users a good amount of information that they search for. For the search engines to achieve the best results, it should first understand what the user is asking for, it should also be capable of differentiating the difference between the sentences containing the same set of words like "the thief threatens the lady" and "the lady threatens the thief". For the most optimal outputs it breaks each data say digital image or media into smaller units called frames and then work on that frames to retrieve the useful information. It uses techniques like object detection, edge detection, sentence tracker and many others. The search engines use different ranking technique in order to rank each set of information on the page that thereafter is provided to the user as a result. Data is classified according to the search technique and through this way they are ranked most relevant to least.

Acknowledgement

We would like to sincerely bring our kind gratitude to Dr. Prateek Jain, Accendere Knowledge Management Services and Ms. Shobha Tyagi, AP, Dept. of Computer Science and Engineering, Faculty of Engineering & Technology, MRIIRS for helping and guiding us in this paper formation.

References

- C. G. Snoek,, & M. Worring, (2008). Concept-Based Video Retrieval. Foundations And Trends In Information Retrieval, 2(4), 215-322. [2]
 C. Snoek et al., "The Mediamill TRECVID 2009 Semantic Video Search Engine." TRECVID Workshop(2009).
- S. Lee, and C. D. Yoo, "Robust Video Fingerprinting For Content-Based Video Identification." IEEE Transactions On Circuits And Systems For Video Technology, Vol. 18. no. 7, pp. 983-988, Jul. 2008.
- Y. Talmi, ed. Multichannel Image Detectors. American Chemical Society, Jun. 1979.
- S.I. Hajeer, R.M. Ismail, N.L. Badr, & M.F. Tolba, (2016). A New Efficient Approach for Multi-Language Search Engines And Formation Retrieval Systems. Asian Journal Of Information Technology, 15(22), 4617-4625.
- M. S. Seidenberg, and J. L. McClelland, "A Distributed, Developmental Model Of Word Recognition And Naming." Psychological Review, Vol. 96. no. 4, p. 523, Oct. 1989.
- S.W. Smoliar, and H. Zhang, "Content Based Video Indexing And Retrieval." IEEE Multimedia, Vol. 1. no. 2, pp. 62-72, 1994.
- F. Idris and S. Panchanathan, "Review Of Image And Video Indexing Techniques." Journal Of Visual Communication And Image Representation, Vol. 8. No. 2, pp. 146-166, Jun. 1997.
- T. Wang, et al., "Person Re-Identification By Video

Ranking." European Conference On Computer Vision, Pp. 688-703, Springer, Cham, 2014.

- Y. Rui, T.S. Huang, and S.F. Chang, "Image Retrieval: Current Techniques, Promising Directions, And Open Issues." Journal Of Visual Communication And Image Representation, Vol. 10. No.1, pp. 39-62, Mar. 1999. C. Snoek, et al. "The Mediamill TRECVID 2009 Semantic Video Search Engine." TRECVID Workshop. 2009.
- J. Sivic, and A. Zisserman. "Video Google: A Text Retrieval Approach To Object Matching In Videos." null. IEEE, p. 1470, Oct. 2003. [13] M. S. Lew et al. "Content-Based Multimedia Information Retrieval: State Of The Art And Challenges." ACM Transactions On Multimedia Computing, Communications, And Applications (TOMM) 2.1 (2006): 1-19.
- M. A. Sadeghi and A. Farhadi, "Recognition Using Visual Phrases," In Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2011, Pp. 1745–1752.
- X. Yu, et al., "Trajectory-Based Ball Detection And Tracking With Applications To Semantic Analysis Of Broadcast Soccer Video." Proceedings Of The Eleventh ACM International Conference On Multimedia. ACM, Pp. 11-20, Nov. 2003.
- M.G. Christel, et al., "Exploiting Multiple Modalities For Interactive Video Retrieval." Acoustics, Speech, And Signal Processing, 2004.
- Proceedings. (ICASSP'04). IEEE International Conference On. IEEE, Vol. 3, Pp. Iii-1032, May. 2004.
- M. Worring, et al., "The Mediamill Semantic Video Search Engine." Acoustics, Speech And Signal Processing, 2007. ICASSP 2007.
- IEEE International Conference On. IEEE, Vol. 4, Pp. IV-1213, Apr. 2007.
- C. Snoek, et al., "A Learned Lexicon-Driven Paradigm For Interactive Video Retrieval." IEEE Transactions On Multimedia, Vol. 9. No. 2, 280-292, Feb. 2007.
- M. Tapaswi, M. Bäuml, and R. Stiefelhagen, "Story-Based Video Retrieval In TV Series Using Plot Synopses." Proceedings Of International Conference On Multimedia Retrieval, ACM, P. 137, Apr. 2014.[20] D. Lin, et al., "Visual Semantic Search: Retrieving Videos Via Complex Textual Queries." Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition, Pp. 2657-2664, 2014.

- R. Kiros, R. Salakhutdinov, and R. Zemel. "Multimodal Neural Language Models." Proceedings Of The 31st International Conference On Machine Learning (ICML-14). Pp. 595–603, 2014.
- A.R. Coden, S.W. Mak, & So, E. C., U.S. Patent No. 5,873,080. Washington, DC: U.S. Patent And Trademark Office, 1999.
- J. C. Monberg, R. Mariani, & S.A. Staab, U.S. Patent No. 6,523,021. Washington, DC: U.S. Patent And Trademark Office, 2003. [24] M.W. Dunn, U.S. Patent No. 5,945,987. Washington, DC: U.S. Patent And Trademark Office, 1999.
- R.M. Krapf, U.S. Patent No. 7,263,709. Washington, DC: U.S. Patent And Trademark Office, 2007.