

A Review of State of the Art Techniques for Multimedia Retrieval

¹S. Bhuvaneshwari, ²R.H.Goudar

ABSTRACT

Multimedia retrieval and its applications ranges from text, still images ,audio, video, animation, graphics and much more. This huge multimedia content can be retrieved based on the text or content. The paper describes the various techniques for retrieving text, image, audio and video. These techniques are based on querying either in text format or content itself. The mechanism which gives results based on the users query by using multimedia content.

Index Terms : Image retrieval, Multimedia, Content based retrieval, Query.

I. INTRODUCTION

Multimedia is an area which deals with computer controlled integration of text, moving and still images, drawing graphics, audio, animation and other media where the data is processed digitally, information can be transmitted and stored. Applications in multimedia contain different media basic characteristics of multimedia system are the use of one kind of media to deliver functionality. They are image, text, audio and video. Text is the most common medium of representing the information. In multimedia text is commonly used for titles, headlines etc. Audio also is included in multimedia which means related with recording, playing etc. audio includes speech, music.

Another important feature in multimedia is video where video means moving pictures with sound. It is the best way to communicate with each other. In multimedia it is used to make the information more presentable and it saves large amount of time. The last feature is image, many multimedia applications have digital image files. These image files use variety of formats and file extensions like png's and jpg's. To provide answers, multimedia information retrieval corresponds best to the users according to the information needs. The information needs which is expressed by the users in the form of queries in information system and later these queries are compared with the database to find relevant information.

II. MULTIMEDIA RETRIEVAL TECHNIQUES

Multimedia retrieval techniques are Image, Audio, Video files can be retrieved using the following techniques.

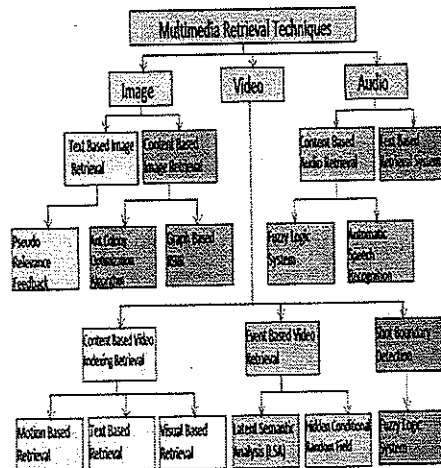


Figure 1. Multimedia retrieval techniques

¹Department of Computer Network Engineering, V T U Belagavi, India. e-mail : bhuvaneshwariscs35@gmail.com

²Department of Computer Network Engineering, V T U Belagavi, India. e-mail : rhgoudar@gmail.com

III. IMAGE

A. Image retrieval Based on Text

The image retrieval based on text which represents the use of images for describing text. For instance the articles in biomedical contains some images which can be interpreted by their connected word. By making use of this text, a group of document is generated to entitle group of images. This type of records are ordered or queried using a text information retrieval based on traditional mechanism, and the documents are searched using based on text queries.

Some of the benefits are there to query and retrieve images based on text using some retrieval techniques. Firstly, this is easily understandable subject, and easily applied for querying and retrieving images based on some related text. Secondly, This is valuable by a reason of some words are distinct documents, image representation can be ordered in data structures which grant for low latency retrieval, finally which grant for image retrieval based on syntactic, the provided mechanism with explanation which enables us to search for images by using the content we required. The image accessing mechanism based on text to be effective, the content of the image must be reflected in the documents that are requested of it [4].

- **Pseudo Relevance feedback (PRF)**

The Pseudo Relevance Feedback (PRF) is the method used for Problem with incomplete annotation can greatly break down retrieval performance. The terms selected by adding feedback to upgrade the user queries from documents which are top ranked in a previous

retrieval. The Pseudo Relevance Feedback (PRF) considers that the target collection provides sufficient feedback information to select useful extension terms. The result display the significant development related to standard PRF methods [8].

B. Image retrieval Based on Content (CBIR)

Image Retrieval Based on Content (CBIR) mechanism describes images as sequence of data elements. Visual features of images have several dimensions like their texture patterns or color. Image retrieval mechanism based on querying a group of images by using a sample image, and it gives the result of the images based on the similarity of visual features with the sample image. When we compared Image Retrieval Based on Content with text-based retrieval mechanism which searches images based on visual similarity of images. For example, group of images are different to each other it gives result by comparing all images with query image [8].

- **Ant Colony Optimization (ACO) Algorithm**

The real ant colonies influences ACO algorithm and functioning way is computational based typical example. The ACO is the member of meta-heuristics class, that has accurate methods used to gain efficient results to difficult issues in a small period of processing duration.

Computational typical example is depends on the examination formed by ethnologists through ants to interact information about smallest way to food through trails pheromone. This process is identified as a feedback loop, due to this chances

of selecting a path increases by number of ants before choosing the similar path. Thus this technique is the reason for the algorithms of the ACO [1].

- **Graph-Based RBIR**

The feedback scheme based on region relevance is combined as a single process of retrieval, based on user by updating importance of query image regions, which leads to development of further performance. By using graph structure and distribution of fuzzy coarse to represent every image, the process of retrieval was implemented by calculating the graph similarity [9].

IV. VIDEO

A. Content Based Video Indexing Retrieval (CBVIR)

The CBVIR which provides the analysis of video by indexing and retrieving. It is flexible and more powerful for video data analysis. CBVIR has several techniques connected with the video indexing based on content and retrieval for video analysis, there is more requirements of index value store and visual information of video is retrieved from the database. The management of the CBVIR most in modeling and querying techniques based on the three abstraction levels the CBVIR can be indexed and retrieved [3].

- 1) **Raw data** : Most low level of abstraction, raw pixels are aggregate to form objects. Comparison is done on the basis by measuring the similarities like correlation coefficient and the Euclidean distance between objects.

- 2) **Feature** : This distinguishes the characteristics of images example shape descriptor, color histogram.

- 3) **Semantic** : The features are collected into meaningful objects at the maximum abstraction level querying is processed out on different entities, with well defined properties [3].

- **Visual based query**

When the user selects one image which should be similar to the one he is searching for. The mechanism gives the result of the images until it matches to the most similar image in decreasing order to the given example of image. So this type of mechanism provides to the user to select number of example so that the query combines all similar features in these group of images. The proposition of the mechanism makes use of the evaluation of search follows in an online learning based on the user's prediction. Positive And negative feedback rates are used as equal NOT and AND operators [3].

- **Motion based query**

It is one of the prime feature of video search engine because this motion query represents temporal info of video content. The complex activity here is the formulation of such queries. These queries can be viewed as counter initiatives which means user needs to specify a motion in still fashion. The task of formulating this query can be made in par with representing motion information in the indexing task [3].

- **Textual based query**

It is the most advantageous approach for querying the video in video database. It has more interface and there are more robust video clips contains hidden clues. Textual data and audio in videos can be extremely large use in indexing. Text information is obtained as captions on the printed / handwritten video or text in the scene. We can identify text content in the video clips, it can be completely used for differentiating the video clips [3].

B. Event Based Video Retrieval

The word event is called shots where remaining shots are called as non-event. By using the event retrieval mechanism event shots are retrieved. There are two types of shots retrieval they are QBE (query-by-example) and (QBK) query-by-keyword approaches. QBK means using keywords user query for event for retrieving and events or shots they are retrieved by comparing with the keywords. User query sample of shots in terms of features they are compared with keywords. The event in QBE is retrieved by using the feature presents in given sample events. This removes the uncertainty linked with syntactic content present in the query-by-keyword method. The QBE method gives the benefits of not requiring predefined retrieval models [7].

- **Latent Semantic Analysis Model (LSA)**

The word's in visual features are associated latent variables, by initiating these latent variables there co-occurrence of formation can be captured in other way. Standard LSA when it is used in image

retrieval the system will improve automatic searching of image. Our method which automatically search recent images are added into the system and those images are retrieved by using image or text queried by the user [6].

- **Hidden Condition Random field (HCRF)**

It has set of hidden states with probabilistic discriminative classifier. Hidden states can be handled by weakly supervised settings. Methods for training video formation, parameter initialization and fusion of different HCRFs are used to establish potential of HCRFs [10].

C. Shot boundary detection

The data which is compressed can be processed in the MPEG and is adorable by using shot boundary detection. The proposed method uses the DCT coefficients of sequences in MPEG. DC images contains transform of every block which results the average value for every pixel used for detection of shot boundary. The obtained solution in a important data and reducing the processing time, to find shot boundary in DC images applying the difference between the pixel metrics and combination of histogram [2].

- **Fuzzy Logic System (FLS)**

The shot boundary detection which make use of FLS for the following reasons i) The FLS method gives positive results, which is far better than single standard of measurement which can be achieved. II) The guidelines can be easily changed without retaining the system. III) The method FLS is computationally inexpensive for development.

Fuzzy systems can be developed by using an expert system the rule based knowledge and membership functions of the process can be generated by the system developer [2].

V. AUDIO

A. Content Based Audio Retrieval (CBAR)

In Audio retrieval attributes are joined and audio data annotations are done manually and it is the conventional method for audio retrieval. The amount of data increases, when the retrieval of information is required is positively similar to huge amount information collected. Some classical properties for different users may even have different context. Hence we require innovative methods for CBAR [5].

• Fuzzy Logic System

When retrieving audio files from huge databases based on user interest it is critical for developing web search engines for multimedia. Based on content audio retrieval technology to differentiate and retrieve clips of audio by using fuzzy character of human perception audio, due to fuzzy character which is automatic according to human perception of audio. There are two important features selected from different feature extraction are used as input to a composed interface of fuzzy system (FIS) [5].

• Automatic Speech Recognition (ASR)

The ASR technique is based on the Audio Retrieval Framework (ARF). The speech related method makes use of the results from speech

processing technologies. ASR gives the important solution to the speech content; compare to speech alone the universe of possible audio is much longer. It is very useful for the users to detect audio files either from huge database of recordings [5].

B. Text Based Retrieval System (TBRS)

The annotation for music based on text-based classical retrieval may be used for finding a audio or music using text documents in a database. which contains the words of a group of a songs collected. In such a situation the retrieval results of audio are visible to a user would contains text paths interrelated to results within the words of the query terms. The retrieval results in the music domain are most common given by exotically playing back parts of an audio recording that consists query terms, while a musical results are the most common form for displaying the query results visually [11].

VI. CONCLUSION

In this paper the first step in internet towards multimedia search engine we have proposed the fuzzy interface system and content based for audio retrieval and classification. The CBIR systems can be used for all preliminary retrieved images so that they can be labeled according to the users expectations into many groups. the methods of user annotations involves several challenges, hence indexing of visual information and video shot annotation is of great demand.

REFERENCES

- [1] Kwang-Kyu Seo, "An Ant Colony Optimization Algorithm Based Image Classification Method for Content-Based Image Retrieval in Cloud Computing Environment", Springer-Verlag Berlin Heidelberg 2012.
- [2] Hong Lu1, Xiangyang Xue1, and Yap-Peng Tan2, "Content-Based Image and Video Indexing and Retrieval" Springer-Verlag Berlin Heidelberg 2007.
- [3] R. Priya, T. N. Shanmugam, "A comprehensive review of significant researches on content based indexing and retrieval of visual information", Higher Education Press and Springer-Verlag Berlin Heidelberg 2013.
- [4] Matthew S. Simpson, Dina Demner-Fushman, "Multimodal biomedical image indexing and retrieval using descriptive text and global feature mapping", Springer Science+Business Media New York (outside the USA) 2013.
- [5] M. Liu, C. Wan, L. Wang, "Content-based audio classification and retrieval using a fuzzy logic system: towards multimedia search engines".
- [6] Nattachai Watcharapinchai1, Supavadee Aramvith1, and Supakorn Siddhichai2 "Two-Probabilistic Latent Semantic Model for Image Annotation and Retrieval" Springer-Verlag Berlin Heidelberg 2011.
- [7] Kimiaki Shirahama · Yuta Matsuoka · Kuniaki Uehara "Event retrieval in video archives using rough set theory and partially supervised learning" The Author(s) 2011. This article is published with open access at Springerlink.com.
- [8] Jinming Min and Gareth J.F. Jones "External Query Reformulation for Text-Based Image Retrieval" Springer-Verlag Berlin Heidelberg 2011.
- [9] Walid Barhoumi, Abir Gallas, and Ezzeddine Zagrouba "Effective Region-based Relevance Feedback for Interactive Content-based Image Retrieval" Springer-Verlag Berlin Heidelberg 2009.
- [10] Kimiaki Shirahama, Marcin Grzegorzec, Kuniaki Uehara "Weakly supervised detection of video events using hidden conditional random fields" Springer-Verlag London 2014.
- [11] Hao Xue, Like Xue, and Feng Su "Multimodal Music Mood Classification by Fusion of Audio and Lyrics" Springer International Publishing Switzerland 2015.

Authors Biography



S. Bhuvaneshwari, M.Tech scholar, Department of Computer Networking Engineering, Center for P.G. Studies, VTU Belagavi. Her area of interest is image processing.