**IOSR Journal of Engineering (IOSRJEN)** **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 13-18

# Implementation of Effective Re-Ranking Approach Using Multimedia Search Engine

Kamatkar Vinalini Vinayak[1], Prof. Jayant Adhikari[2], Prof.Rajesh Babu[3]
[1]*M.Tech Scholar, Department of Computer Science and Engineering Tulsiramji Gaikwad-Patil College of Engineering and Technology Nagpur, Maharashtra, India*
[2,3] *Dept. of Computer Science and Engineering Tulsiramji Gaikwad-Patil College of Engineering and Technology Nagpur, Maharashtra, India*

***Abstract:*** *Recently the detonating growth and comprehensive accessibility of association share media content on web have induce rise of research action in multimedia search. Generally while seeking result not actually coordinates with the search results. Hence the growing images on internet requires the best image retrieval technique that can enhance the image retrieval accuracy. Re-ranking is a viable technique to improve the search results of internet based multimedia search. This is generally adopted by business web crawlers, for example Google. The proposed Re-ranking approach is fit to work with all media types: video, picture, and audio. Techniques that apply text search methods for multimedia search have achieved limited success as they completely ignore visual content as a ranking. The crawler are generally base on content and compelled as the user look by keywards which results into vulnerability among multi media. Because of which loud or irrelevant images or video are available as results. The aim of multimedia search re-ranking is to reorder retrieved content to get optimal rank list. For that group of descriptors are utilized with weight and weight are assigned to it dynamically for acquire accurate multimedia files. Here we discuss various techniques for web multimedia re-ranking and propose new re-ranking technique to achieve the accurate query result and result demonstrates that it retrieves most significant files to the top. Here we utilize* ***Optical Character Recognition (****OCR) and ASR i.e. combination of OCR and ASR(Hybrid Algorithm) to get text content from image and video.*

***Index Terms:*** *Re-ranking, Multimedia Retrieval,OCR, ASR, Audio-Video Feature Extraction.*

## I. Introduction

From past few years, web has been spread widely all over the globe, since it has huge multimedia database over the internet. Searching the exact multimedia that is image, video or audio from such a large database is a very complicated task. Internet scale search engines utilizes mainly two approaches. First one text-based image search. Many commercial internet scale multimedia search engines utilize this approach. They use only keywords as search queries. Users generally fire query on crawler in the hope of getting a certain type of multimedia.

The multimedia Re-ranking is an efficient technique to enhance the results of internet based image, video or audio search. It has been developed by existing industrial search engines like Google, Bing and so many. For an inserted query keyword, crawler re-ranks the group of multimedia depends on the fired command. Then the user select a particular multimedia from the group, the rest of the multimedia are re-ranked based on the user chosen multimedia.

It is famous in all kinds of crawlers. But it gives ambiguities in result. For instance user has given input query as ‘colors, so as the entered query is not specific system can retrieve images like different colors, colors logo, colors TV or videos on colors TV knowledge about query keyword else he won’t get useful images. The semantic meaning of command keyword may be different than intended. The search engine gives additional text keyword suggestion when user enters the command. It’s beneficial but it may possible that user may get diverted from its way. Colors TV shows, photos etc.

Multimedia search is a broadly growing feature of well-known crawlers such as ‘Google’, `Yahoo', `Bing', etc. For a given text query, the search engine has to search for millions of images, video and audio for retrieving the significant multimedia as early as possible. Generally the search engines are depends on text meta-data as keywords, tags or text descriptions near to the multimedia. As the meta-data do not constantly be in associate to the visual term of the multimedia, the retrieved multimedia is normally collected with unwanted irrelevant multimedia.

Web multimedia re-ranking is a process in which images, audio, video are retrieved, arranged as per their features and user need. Multimedia search has become an increasingly vital research topic to facilitate access to the fastly growing collection of multimedia on the internet and exploit their benefit for the users. Web scale multimedia search engines mostly utilize keyword as queries and depend on the surrounding text to search the multimedia. They generally operates in two main steps: the offline index generation and the online index serving. Meaningful and efficient multimedia retrieval is a challenging task for efficient web search. By utilizing existing methods uncertainty is occurred in multimedia retrieval. Sometimes accurately describing the visual content of target multimedia using keyword only is difficult for user. Due to lack of query and visual features multimedia obtained are less relevant as per user requirements.

To locate the problems of visual search approaches, multimedia search re-ranking has received increasing heed in recent years. It is defined as the reordering of visual documents based on the information manifested in the initial search results or a knowledge base to improve the search performance. This information consists of multimodal cues that can be the knowledge or specific patterns mined from the initial ranked list, query examples or any available knowledge. Re-ranking approach specially find the user activities and used to enhance the percentage of the retrieved relevant files. We believe that if a user did the actions on any of query files known as Target is related to the required ones. Then the Target file will be examined and the query results will be rearranged based on their similarity with the Target.

Optical Character Recognition (OCR) utilizes optical scanner or computer applications to convert the non-editable text to editable text. OCR generally used to get data on image and translate it into ASCII code. There are number of software's available to extract frame from video.

Key frame extraction method generally utilized to protect video copyright. This method is depends on frame difference with low level features, consisting color feature and structure feature. To extract correct key frames two stage strategy is utilized and to cover the content for the whole video sequence. Depends on color characteristic difference among adjacent frames from original sequence is taken known as alternative sequence. Then by examining that difference final key sequence is obtained. Then in order to assure performance of key frame extraction optimization step included based on number of final key frames.

Automated Speech Recognition (ASR) is a technology to convert speech or video words into written text. ASR is used for getting text from video. For that they utilize device. It work same as OCR but OCR get text on image and ASR get text from speak content.

A multimedia search engine generally operates in following steps: Offline and Online operation

### *A. Offline Operations*

When a unexisted file is added to the database, before saving the file first it is preprocessed and extracted features from them.

#### *1) Pre-processing operations:*

Preprocessing is utilized to discard unused parts i.e noisy part of multimedia since multimedia files are most frequently captures and edited by doing so multimedia files becomes noisy. It will affect the quality of multimedia retrieval.

#### *a) Preprocessing operation on image files*

Preprocessing adopt linear, nonlinear like median filter, weighted median filter and fuzzy methods to remove noise in image files. Linear filters may diminish image details, Fuzzy filters provide promising result and median filter detect to be efficient in storing image detail at cheap cost

#### *b) Preprocessing operation of Video files*

Video structure contains various continuous frames merged with the audio data. Sometimes it found that it consist some meaningless frames like totally black, white, faded frames. This makes the process of selecting valuable features for representing video files complicate and time consuming. Static video summarization method utilized as it needs less time than dynamic method for filter and remove the irrelevant and padded video frames.

#### *c) Pre-processing operation of Audio files*

Audio file formats consist bit rate, sampling rate, and number of channels and that will indicate that two similar files are distinct. To resolve this problem, the audio files will be decoded with Pulse Code Modulation format (PCM).

***2) File features Extraction***
Feature extraction from image, video and audio files is described in this section:

***a) Feature extraction process from image***
To extract features from image media Scalable Color Descriptor (SCD), Edge Histogram Descriptor (EHD), and Joint Composite Descriptor (JCD) utilized.

***b) Feature extraction process from video***
Fingerprinting is utilized for processing of large database and decreasing cost of computational file.

***c) Feature extraction process from audio***
Fingerprints of the audio files consist of :i] Spectrogram construction: After preprocessing the signal with 5512 Hz PCM, the audio spectrogram, which shows how the spectral density of a signal varies in time will be constructed.ii]Band filtering: Generally human ears can recognize the frequencies in the range 20 Hz-2000 Hz. In this spectrogram is utilized to get the range which human eye can sense.

***B. Online Operations***

Here user behavior will be supervised. When the user download, copy, or spending more than a number of seconds (N) with a Target file, the results of query will be re-ranked by applying following: i] Estimated distance among the Target file (T) and the other files (Y) by using the Euclidean formula. ii] A weight assigned to each descriptor. This is based on the similarity of the descriptor vector of the Y and T files. In case of images and videos, the descriptor capability to differentiate between the files will be assigned a higher weight than the others, while the opposite is applied in the case of audios. iii] Normalize and compute the rank of the file Y.

The section I explains the Introduction of multimedia re-ranking techniques. Section II presents the literature review of existing systems and Section III present proposed system Section IV presents experimental analysis of proposed system. Section V concludes our proposed system. While at the end list of references paper are presented.

## II. Literature Review

Zhu et al. proposes a new re-ranking scheme [1] and demonstrate performance results experimentally for web image retrieval with integrated query. Also they worked on cross-modal association rule that was designed for associating one keyword with several visual feature clusters in web image retrieval. He developed an automatic re-ranking process online to integrate the keyword and visual features for web image retrieval depends on the cross-modal association rule, and gives experimental test. The experiment is carried out on web image retrieval system named VAST (VisuAl & SemanTic image search).

Chen et al proposed a semi-supervised ranking aggregation method [2], in which the order of several item pairs are stated as side information. The main thing is to learn a ranking function depends on the ordering agreement of different rankers. The ranking scores assigned by this ranking function on the labeled data are consistent with the given pairwise order constraints while the ranking scores on the unlabeled data

Due to vast growing online content, automated search engine programs are utilized to search and divide millions of webpages and show only the most relevant pages for the query search given as an input by the user. Crawler employ combination of automated algorithms, manually edited directories and advertisements to establish results for user's queries. He proposed enhanced version of the standard PageRank algorithm [3] by utilizing weight of in-linked web pages. Than evenly dividing the weight of an in-linked webpage, their technique distributes it to all the out linked pages based on their popularity. Also invented augmented page ranking algorithm as WIL (Weightage In-Link) PageRank algorithm. WIL utilizes the weights of in-linked webpages to estimate a new score of every individual webpage known as WIL-score. Then according to this WIL-score the webpages can be ranked.

A novel Ranking Preserving Hashing (RPH) approach [4] proposed by Lu, M., Huang et al. to optimize a popular ranking measure, Normalized Discounted Cumulative Gain(NDCG), to obtain effective hashing codes with high ranking accuracy. The main difficulty in the direct optimization of NDCG measure is that it based on the ranking order of data examples, which forms a non-convex non-smooth optimization issue.

For online image search with re-ranking a real time image search engine is developed. The proposed Adaptive Similarity is motivated by the idea that a user always has a specific intention when submitting a query image. But it was difficult for the eight weighting schemes to cover the large variety of all the web images. It was also likely for a query image to be classified to a wrong category [5].

Retrieval is based on the query-by-example paradigm, the user gives a query image, for which a semantic multinomial is computed and matched to those in the database. QBSE [6] produces retrieval systems that are more accurate than existed one. Introduction to a Human Computer Interaction approach to CBIR based on relevance feedback. It is not like the computer centric approach where the user has to precisely decompose his information need into different feature representations and precisely specify all the weights associated with them the proposed interactive approach allows the user to submit a coarse initial query and continuously refine his information need via relevance feedback. For web scale user's feedback has to be constrained. This approach greatly decreases the user's effort of composing a query and captures the user's information need more evenly[7].

W. Y. Ma et al. proposed NeTra, which is a prototype image retrieval system. It utilizes color, shape, texture and spatial location information in fragmented image section for searching and extracts similar section from the database. The search depends on object or region is permitted in this system and the quality of image retrieval is also enhanced when images include many complex objects [8].

Most of Pseudo-Relevance feedback techniques limit users effort by widening query image with maximum visually similar images. R. Yan et al. present a concept to give user approximate images in just a one click. Semantic gap between query image and other visual inconsistent images results into poor performance. In this, top N images which mainly visually match with the query image are considered as extended positive examples for obtaining a resemblance metric. But the top N images are not essentially semantically related to the query image, thus the obtained resemblance metric may not always show the semantic relevance and may even deteriorate re-ranking performance [9].

J. Cui et al [10] classify query images into eight pre-identified intention classes and different types of query images are given different feature weighs. But the large variety of all the web images was complicated to cover up by the eight weighting schemes. In this, a query image was to be categorized to a wrong class.

Cai et al. [11] prescribed matching the images in semantic spaces and re-ranking them with attributes or reference classes which were manually defined and learned from training examples which were manually labeled. They consider that there was one main semantic class for a query keyword. Re-ranking of images is done by utilizing this main category with visual and textual features. Still it is tough and not efficient to learn a universal visual semantic space while expressing highly varied images from the web.

## III. System Architecture

### *C. System Architecture*

Here in Fig.1 Multimedia (Image, Audio and Video) dataset is taken as an input after that preprocessing is applied on that to extract features. Among the extracted features essential features are selected for further processing after that similarity between user query and dataset files is calculated using algorithm and to improve result accuracy re-ranking of targeted output is performed.

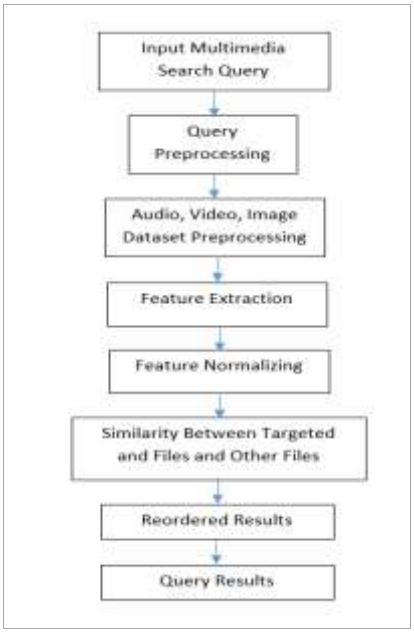


**Fig 1.** System Architecture

### *D. Proposed System*

Proposed Media (Audio, Video, Images) Retrieval Systems executed on four primary modules such as features extraction, media retrieval utilizing queryfeature extraction with similarity measures, keywords extraction using ASR techniques, and at the end OCR extraction is performed to get HOG, image color and other important featuresas appeared in figure 1.Firstly, a user transfers or gives a question as input to the Personalized framework. System will partition the casings into video and does choice procedure of pertinent edges into all edges. At the same time ASR framework will process on video info and concentrate the catchphrases by ASR procedure. After edge division and choice, perform OCR and concentrate the HOG, OCR content and Gabor Filter from chosen outlines and furthermore separate the Color, Texture and Edge identifier. A similar procedure of ASR, Frame division, OCR and picture handling is done on recordings put away on database or folders. After pre-processing framework will look for comparability in keywordsand highlights of inquiry media metadata and all media which are put away in database or folders. Our framework extricates the most coordinating OCR content, ASR content and keywords and includes and creates important last media results.

## IV. Result And Discussions

### *E. Experimental Setup*

All the experimental cases are implemented in Java in congestion with NetBeans tools and. algorithms and strategies, and the competing OCR and ASRapproach along with various feature extraction technique, and run in environment with System having configuration of Intel Core i5-6200U, 2.30 GHz Windows 10 (64 bit) machine with 8GB of RAM

### *F. Result Analysis*

Table1 shows the precision and recall values for video retrieval with combination of OCR and ASR (Hybrid).Our proposed approach can do some improvement to existing approach by doing combination of OCR and ASR. As the value of precision and recall of Figure 2 For personalized video results sequence should change according to that user query interest.

**Table 1.** System measure analysis for multiple search query

| **Search Query Measures** | **Q1** | **Q2** | **Q3** |
|---|---|---|---|
| Precision | 0.91 | 0.87 | 0.85 |
| Recall | 0.87 | 0.81 | 0.79 |
| F-measure | 0.87 | 0.84 | 0.84 |

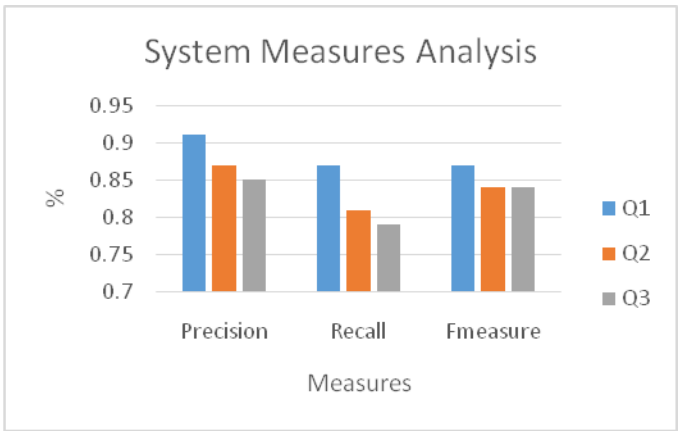


**Fig. 2** System measure comparison for hybrid (OCR and ASR) algorithm

## V. Conclusion

Here we discussed Re-ranking techniques utilized in multimedia search engine like video, image and audio for ranking the multimedia to get accurate results. The proposed re-ranking system gives better results of web-scale multimedia re-ranking than the existing system and also considerably performs better in both the accuracy and efficiency of the re-ranking method. Multimedia search re-ranking is utilized to reorder retrieved contentfor getting optimal search rank list. Re-ranking utilizedto get the accurate query result. Here we utilized concept of OCR and ASR. Optical Character Recognition referred as OCR. It utilized to get text or non-editable data on image, it select text on image and convert it to editable text format. Automated Speech Recognition

(ASR) utilized to recognize speech data and get it in text form. By utilizing ASR and OCR we can get text data then features extracted based on that to get accurate query result.

## References


[1]. Zhu, Y., Xiong, N., Park, J., and He, R., "A Web Image Retrieval Reranking Scheme with Cross-Modal Association Rules", International Symposium on Ubiquitous Multimedia Computing, Issue 13, Pages 83 - 86, 2008.

[2]. Chen, S., Wang, F., Song, Y., and Zhang, C., "Semi-supervised ranking aggregation", Information Processing & Management, Volume 47, Issue 3, Pages 415-425, 2011.

[3]. Singhal, R., & Srivastava, S. R. , "Enhancing the page ranking for search engine optimization based on weightage of in-linked web pages." Recent Advances and Innovations in Engineering (ICRAIE), 2016 International Conference on. IEEE, 2016.

[4]. Lu, M., Huang, Y., Xie, M., and Liu, J., "Rank hash similarity for fast similarity search", Information Processing & Management, Volume 49, Issue 1, Pages 158-168,2013.

[5]. J.Cui, F. Wen, et.al, "Real time Google and live image search reranking", The 16th ACM international conference on Multimedia, Pages 729-732, 2008.

[6]. X. Tang, K.Liu, J. Cui, et.a, "IntentSearch: Capturing User Intention for One-Click Internet Image Search", IEEE Transactions On Pattern Analysis and Machine Intelligence Vol. 34, No.7 pages 1342 – 1353, July 2012.

[7]. Y. Rui, T. S. Huang, M. Ortega, et.al, "Relevance feedback: a power tool for interactive Content-based image retrieval", IEEE Transactions On Circuits and Systems for Video Technology, 1998.

[8]. W. Y. Ma and B. S. Manjunath, "A toolbox for navigating large image databases, multimedia system," 3(7), 1999, 184-198.

[9]. R. Yan, E. Hauptmann, and R. Jin, "Multimedia Search with Pseudo-Relevance Feedback," in Proc. Int. Conf. Image and Video Retrieval, 2003.

[10]. J. Cui, F. Wen, and X. Tang, "Real Time Google and Live Image Search Re-Ranking," in Proc. 16th ACM Int. Conf. Multimedia, 2008.

[11]. J. Cai, Z. Zha, W. Zhou, and Q. Tian, "Attribute-Assisted Reranking for Web Image Retrieval," in Proc. 20th ACM Int. Conf. Multimedia, 2012.