Video Search Engine Optimization Technique Using Keyword and Feature Analysis

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering

in

Software Engineering

Submitted By

Krishna Dinesh Choudhari

(801331007)

Under the supervision of

Mr. Vinod Kumar Bhalla

Assistant Professor

(CSED)

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004

July 2015

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled, "*Video Search Engine Optimization Technique Using Keyword and Feature Analysis*", in partial fulfillment of the requirements for the award of degree of Master of Engineering in *Software engineering* submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of *Mr. Vinod Kumar Bhalla* and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

(Krishna Dinesh Choudhari)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

Mr. Vinod Kumar Bhalla

Assistant Professor

CSED, Department

Countersigned by

(Dr. Deepak Garg)

Head

Computer Science and Engineering Department

Thapar University

Patiala

(Dr. S. S. Bhatia)

Dean (Academic Affairs)

Thapar University

Patiala

Acknowledgement

First of all, I am extremely thankful to my respective guide ***Mr. Vinod Kumar Bhalla*** of Thapar University for his valuable guidance, advice, motivation, encouragement, moral support, sincere effort and positive attitude with which he solved my queries and provide delightful ambiance for learning, exploring and making this thesis possible. It has been a great pleaser and experience to work under his sanctuary.

Dr. Deepak Garg is one of the best teachers that I have had in my life. He sets high standards for his students and he encourages and guides them to meet those standards.

I would like to thank my family members who are dearest and precious to me for their love, encouragement, blessings and support in all respects. Most importantly, none of this would have been possible without the love and patience of my family. My immediate family, to whom this dissertation is dedicated to, has been a constant source of love, concern, support and strength all these years.

I am also thankful to the entire faculty and staff members of Computer Science and Engineering Department for their direct-indirect help, cooperation, love and affection, which made my stay at Thapar University memorable.

Date: July, 2015
Place: Thapar University, Patiala

(Krishna Dinesh Choudhari)

Abstract

Growth of internet touched upon every sphere of life. Business is no exception to it. More and more companies and individuals are bringing their business online. Now a days videos are used as a tool to advertise and promote the business. Enterprises upload relevant videos on such promotional sites so that people can extract the most relevant video content. Users tend to choose the top ranked and the most viewed videos irrespective of their relevance. So, the key to provide suitable content is Video Search Engine Optimization (VSEO). This research work proposes a method to optimize the video rank by exploiting video search engine's searching strategy that will eventually lead to increase in number of views of videos with higher ranks and thus promoting corresponding website for every visit. To promote white hat SEO, a technique is suggested that uses keyword tags in title, description and transcript. Ranks of videos are analyzed before and after VSEO. The main idea of the proposed strategy is to select appropriate keyword tags based on navigational search queries, transactional search queries and informational search queries. Experiments are performed on two major video tube sites to establish the credibility of proposed technique. Prominent results are achieved by applying proposed methodology on various case studies. Suggestion list is generated for common users to get higher ranking by following the technique presented in research work.

Table of Content

List of Figures

List of Tables

Chapter 1
Introduction

There has been a tremendous growth in digital market recently, because of the focus shift to visual content and video broadcasting for marketing. The increase in the bandwidth provided by the telecom's Internet Service Providers (ISPs) has resulted in lag-free billions of hours of videos being watched on various video search websites. Publishing a video on the internet is the cheapest source of advertisement as compared to conventional TV advertising campaign. Therefore, many commercial organizations are emerging with their channels on video search websites where they publish and earn money though the traffic fetched by the video, also it is easy for them to link it with company's main website & other social media. Video search engine enables user to choose and watch videos of same category by recommending other related videos. Moreover, through video search engine user can recommend and share videos with family and friends. These videos become visible to them in their related videos option. In order to increase the view of videos, producer has to make effort to boost the ranking of the video, this process is called video engine optimization. Thus higher the rank in the search result, the more will be the count of visits received, which brings more sales revenue through the website visits. Many believe that delivering lot of related videos will result in top ranking, but it is true only if the content is purely excellent, it will get views by people and thus reinforcing the prominence of the content. The videos that are ranked at top in search results do not only rely on the content but also on content delivery's right format.

In the field of search engine marketing, videos are play the important role in visiting of any website. Internet marketing strategy include video promotions. Mostly videos are search through video search engines. User enters the keyword into the search field and relevant result are given by search engine and pick one video which it's seems that most appropriate. User normally choose those videos that ranked top by search engine. Even if the content of the video is relevant then also video is getting lower ranks in search. This research work poses the video engine optimization technique to improve the visits and view of video as a result apparently improvement website visits.

1.1 Assortment of Search Engine

There are four types of search engine according to their service providing way and information collection methods.

1.1.1 Text search engine:

Text search engine extract any website information from the World Wide Web, create data sets and access the record what is similar to the conditions given by the user and return the result using certain combination and permutation. It consist of Self built WEB database, it has its own web grab mechanism. Indexing and indexer has independent "spider" program or crawler or "robot" program.

1.1.2 Directory Indexing:

Web link list are classified by the catalog. It produces result without inquiring keywords. This type of search engines cannot be called a true search engine also the submission of the pages is also manual.

1.1.3 META Search Engine:

Meta engine uses multiple search engines simultaneously, after processing through multiple search engine, it returns the result to the user. Uses set rules and do proper combination and permutation for producing result [1].

1.1.4 Vertical Search Engine:

It is a demand search and search in the specific field for better user experience. To reduce the hardware and computation cost this type of specific field search engine used. It works on particular user needs and more varied inquiry modes.

Table 1.1: Types of search engines [1]

Search engine types	Example
Text based	Google, Alltheweb,Teoma,Bing,Yahoo
Directory indexing	Youtube,MSN search,Dailymotion
Meta search	Dogpile, Mamma, Metacrawler
Vertical search	Mocavo, Nuroa, Trulia, Yelp

1.2 Difference Between SEO and VSEO

Table 1.2: Difference between SEO and VSEO

Search Engine Optimization	**Video Search Engine Optimization**
Optimization is done for website.	Optimization is done for video only
It has main feature for optimizing is heading tags.	It has main feature for optimization is Title of the video.
Crawler uses Meta description to get idea about website.	Crawler uses video Description to get idea about video.
Website rank is based on number of internal link and external links from other website.	One the factor for deciding video rank is embedded video link and external links from other social networking site.
Properties on which SEO can be done are Image alt tags, Image format, Page size, Image size.	Properties on which VSEO can be done are annotations, transcript, video length, video format and video size.
Trust rank on the basis of number of hits, clicks and links.	Trust rank on the basis of views, rating, comment, description.

1.3 Recommended Optimization Techniques

This section explores methods that are suggested as best approaches for search engines. They may be sets as follows: on-page and off-page, keyword, indexing. The properties of each optimization category will be discussed in the following subsections.

The three important process of search engines are: accessing the web and observing important words (crawling), constructing and maintaining an index of sites keywords and internal links (indexing), and return the result based on reputation and importance to the user's search terms (searching). Near about all the search engine working involves keywords research [2].

1.3.1 Keywords Choice:

One of the most influencing factors in the search result ranking of keywords, because it returns will directly correlate to a video. During following VSEO techniques, choosing keywords and its application, webmaster should choose those keywords which are important and potentially those keywords which belong to business, associates or individual has a customer base associated with it. Through the analysis and comparison,

one can calculate keyword demand. Since, Video search engines records the every words and sentence that is typed into search engine, will be available in the keyword research tools, owner of the website will be able to make decision by gain a general comparison of high-demand keywords within their niche customer base.

As in SEO of website do keyword research for online marketing, for that one has to target audience of video and have to do research like what audience will place the keyword inside the search engine.

There are many tools available which analyses the keyword density. For the website classification, keyword research tools can be used for deploying keyword columns. The existing era there is different there are different tools are present like keyword selector tool, keyword external tool, keyword research tool. The keyword density in keyword analyzer tool is preferred to the ratio of text of a key or phrases appearing. In the present scenario there are some tools refer as competition determination tool - seoscorecard, Top completion examination tool- WebuildPages, Google Adwords, Youtube keyword suggestions or query log analysis [3].

1.4 Indexing:

This section explores techniques to way of accessibility of a video to search engines. To ensure the visibility in search results, video links and content can be search and crawled by search engine is an influential step. The techniques that can be used to improve the probability of videos crawl-ability will be discussed. After that, what tactics should be applying in order to make the help the robot to search the video easily will be discussed [4].

1.4.1 Optimization of Index

There are two approaches that confines new video are get search and indexed. The publisher has to list those videos that they want to get it crawls and index.

1.4.2 On-Page Optimization

In the different stages of optimization on-page optimization is intended to support search engine for the end user style to drive the traffic. It includes the various alterations of

elements on-page. This explores the involvement of alteration of keywords from the earlier stage to the page is to be optimized. Arrangement of the keyword must be place on-page from the human perspective whereas there is an extra burden on preserving the principle of white hat techniques. This section explores strategy to place relevant keyword.

1.4.3 Off-Page Optimization

Link building process is influential fundamental aspect in the making of PageRank. Most of the marketing page owners do this for purpose of obtaining links from other site, because links not only use to search pages and video but also examine ranking and how thoroughly and rapidly a website is crawled.

It is recommended that Facebook and Twitter can play vital role in link-building factors. As of 2011, it was confirmed that both Google and Bing treat shared link on walls and Facebook fan pages as votes for the website being shared. In short social media communication provides best solution for off-page optimization and it improves the PageRank which help to build strong brands [5].

1.5 Need of Video Search Engine Optimization

1) Video boost 40% sales of retailer's site. People who watch video are 1.81 times more likely to buy an item than non-viewer [6].
2) Now a days marketers are using visual content for online advertising 93 % of the time [7].
3) Video apps are increasing more frequently visit of mobile user over mobile web [8].
4) Usage of "video" keyword in subject field of email increase the chances of open rates by 18.5%. It also increase click through rates by 64.8% and reduce the chances of getting unsubscribes by 26% [9].
5) 60.4% audience decide to purchase item after watching video email [10].

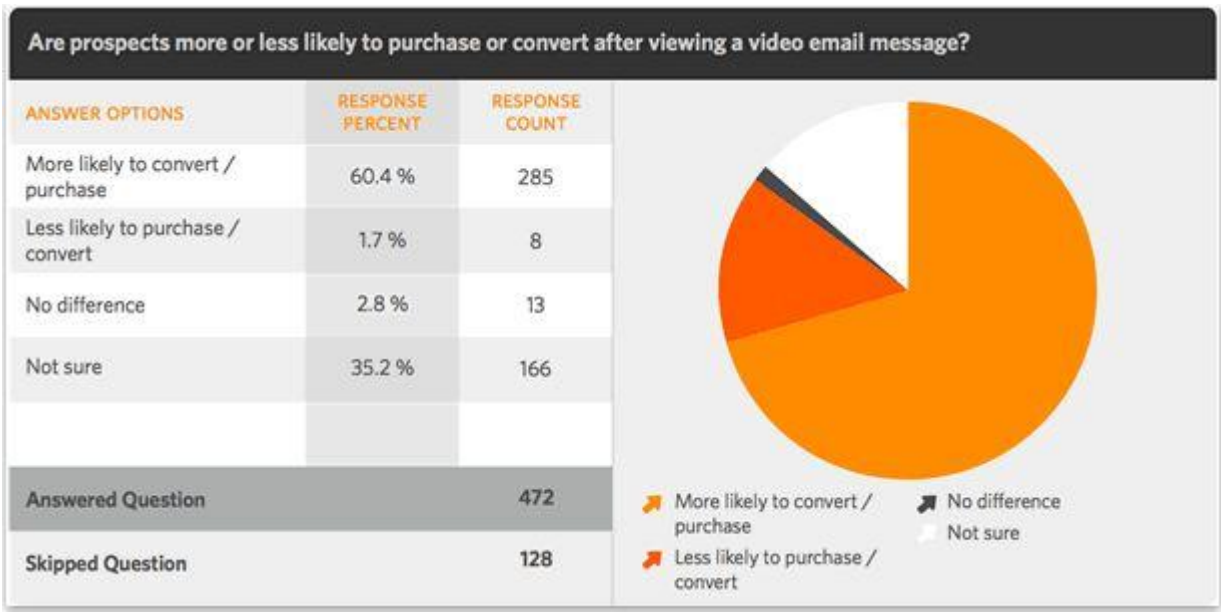


Figure 1.1: Response of Marketing through video email [10].

On analyzing of Figure 1.1 and studying the above mentioned points, it is observed that usage of video search and advertising of various services/products over the internet has been increased exponentially.

1.6 Organization of Thesis

This thesis is organized as follows

Chapter 1: Introduction of Search Engines and its approaches for optimization.

Chapter 2: Earlier work on the Video Search Engine is represented in this chapter.

Chapter 3: What are the various problems faced during Video Search Engine optimization is represented in this chapter.

Chapter 4: The methodology for optimization are represented in this chapter.

Chapter 5: This chapter represents implementation and result of the proposed work.

And the last chapter is devoted to the conclusion.

Chapter 2

Literature Survey

VSEO depends on retrieving videos from database according to query search, therefore, it is necessary to identify how the video search engine model has been changed for satisfying the user needs.

The keyword and video indexing model is used to enhance user experience. Firstly video indexing is done by dividing video lecture into parts called as segment. To analyze long video present in the database, binary search with frame sampling is used. Keyword indexing involves extraction of keyword from video segment. Image is enhanced for the extraction of keyword. Research also analyzed important of video in learning [11].

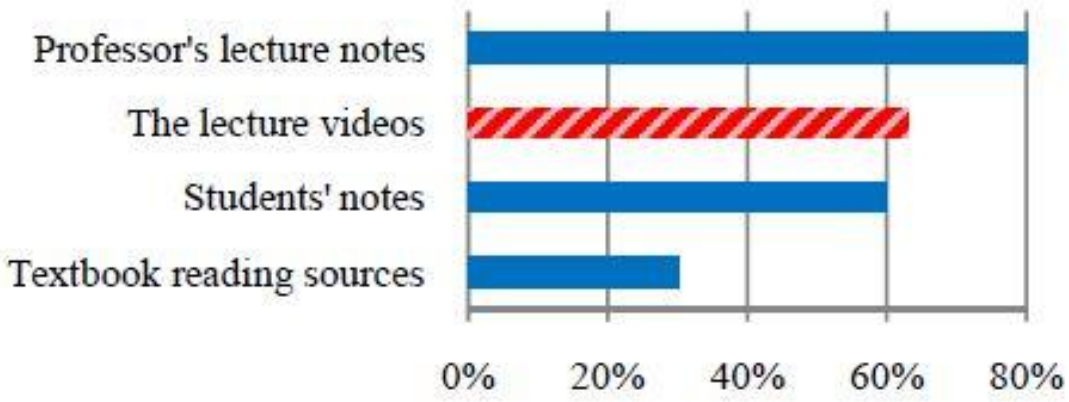


Figure 2.1: Importance of lecture video & other learning resources [11].

Early video search engine retrieving is rely upon color, edge and text. System takes the text as input, after matching query text it returns relevant video. System based on this criteria was time consuming as it saves video in the form of cluster and retrieve through database.

Later on many model came [12,13,14], based on search query criteria and there is only significant change in search filtration criteria. System enables users to choose positive and negative samples from extracted videos. This system filters the query on the choice made by user and fetches a newly created list of videos.

In the study of search engines query log it has been analyze that viewers are usually choose video that in top ranking rather than mostly viewed video. CR reranking which uses cross reference (CR), helps to retrieve video effectively with more accuracy. Model uses initial ranked result and makes its cluster for reranking. Reranking model achieved high accuracy in showing top- rank result.

Literature define in the beginning on VSE retrieving of video is based on surrogate. Video surrogate is used to represent video content. This representation shows extra feature of true content of video. Surrogate can be keyword, photographs, .gif image file that is skimmed video and audio based. Surrogate helps user to find detail about the video. Audio content can be consider as factor to record the keywords. Low speed in processing video and audio content is time consuming [15,16,17].

In later studies it was found that fast forward method (FF), speed is improved as no audio is attached and hence it saves time. Fast forward method is video content summarization technique. This method involves cutting of video into frames. Cutting of video into frames also consume time [16,17,18], so adding complementary metadata like voice attribute will increase processing time.

In order to reduce the time consumed for producing related material, video learning platform uses computer generated surrogate. For better comprehension, keyword is textual surrogate that helps searching the specific subject.

For such an effective comprehension, keyword based video summarization platform (KVSUM) presents a keyword cloud that not only help to support to organize video information but also adds textual surrogate.

2.1 Basic Principle of Video Search Engine

FF and KVSUM are two mechanism for summarization. Average sampling makes FF and KVSUM represent text and visual as separation data and uses the keyword cloud. Figure 2 presents KVSUM system, while figure 2.2 represent FF system. As shown in Figure 2.2, KVSUM consist of five stages, keyword extraction, subtitle recognition, summarization stage, thumbnails creation and keyword cloud creation. Raw visual content or video is input to the KVSUM. It generate output of thumbnails, video frames with thumbnails and transcription.

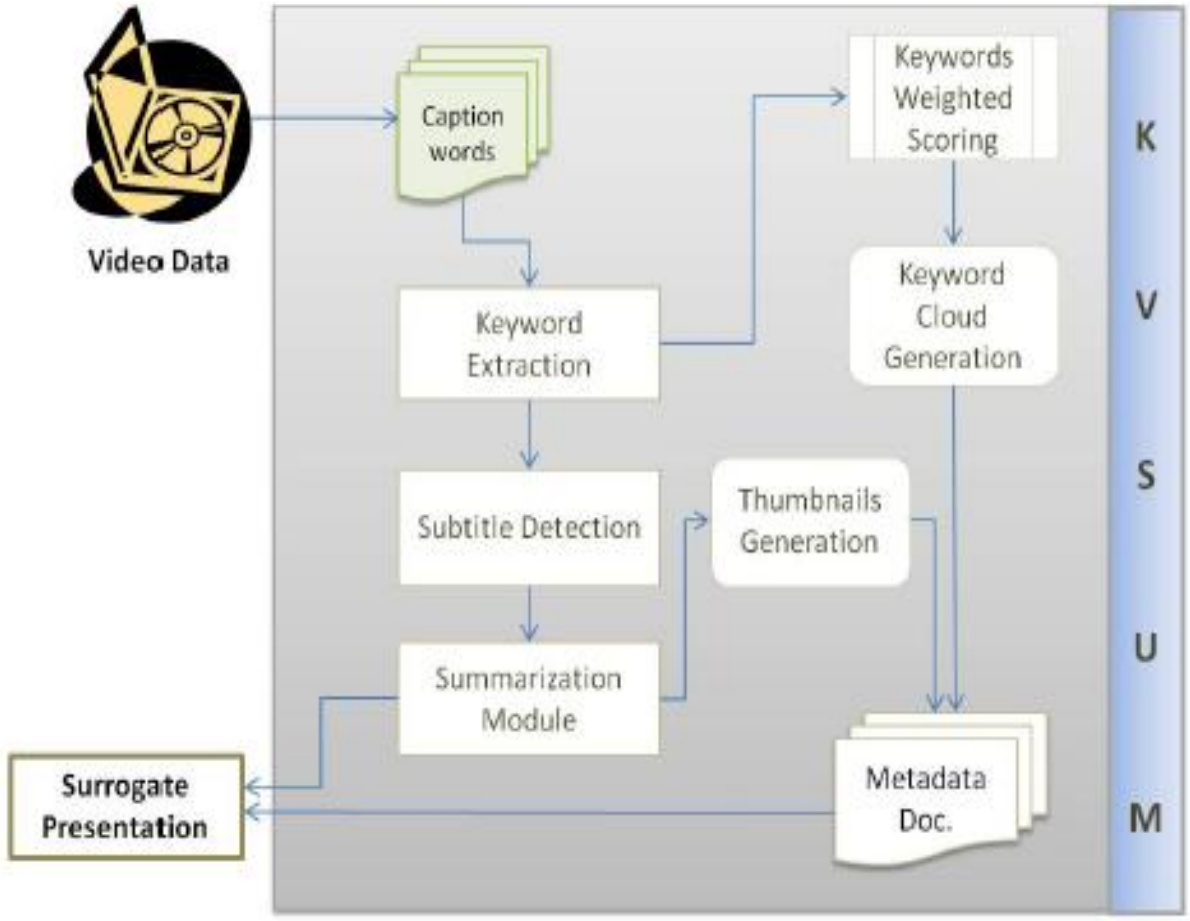


Figure 2.2: Model of keyword based video summarization platform (KVSUM) [19].

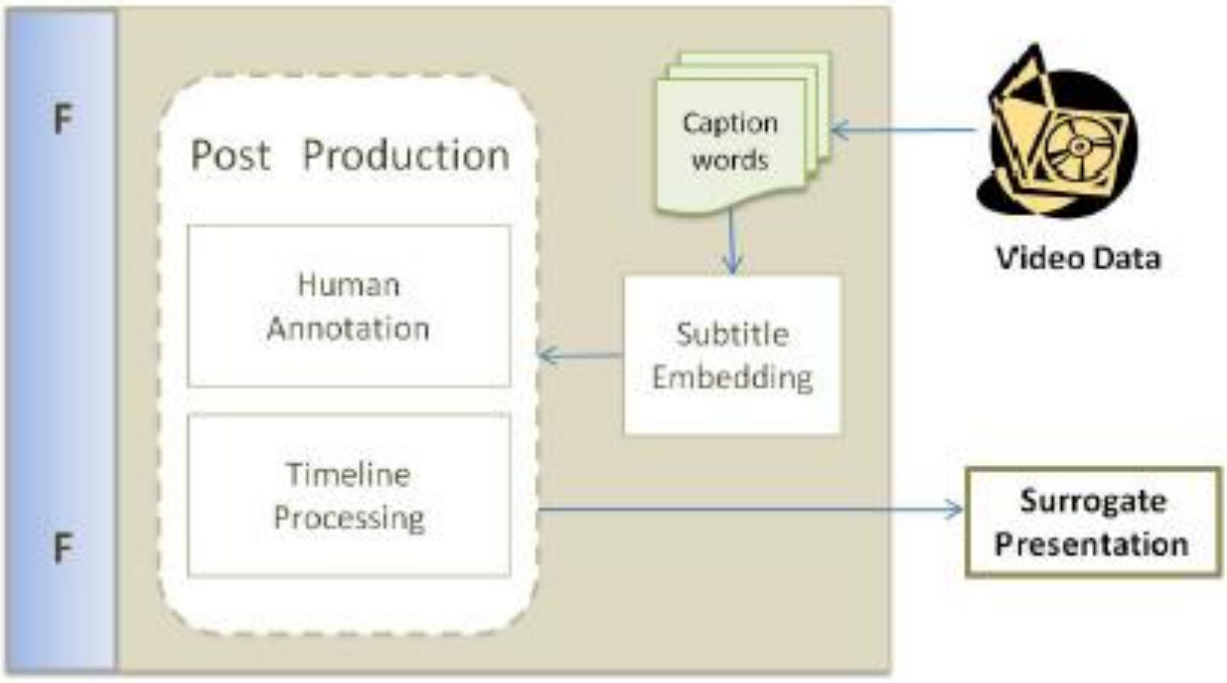


Figure 2.3: Model of Fast Forward (FF) approach [19].

Raw input video are break into three parts: video frame, time-stamps and subtitles. Thumbnail is created by making frame set of the video. To generate summary, subtitle are split to discover the keywords. This extracted keyword takes into consideration while summarize the video. Summarization module creates the different size of video summary by limiting time.

After creation of keyword, as per the presented weighting scheme, keyword cloud creation module places the keyword into the classic tag cloud. At the last stage, integration takes place of generated keyword, summary generated by KVSUM and generated video frames. To access the video summery, users can fetch the keyword data from cloud [20].

The FF is related to efficient compression of video. The FF work on time constraint, it plays the video faster than normal speed for example if the time of video is half then it plays video two time faster. The FF bound video to the time, more will be tightly bound more will the playing speed. The human annotator is set the caption to give support to the learners. This is the post processing, which is a bit time consuming.

2.2 Keyword Weighting Scheme

To recognize the keyword characters, processing for identifying character as word is depend upon assessor variety score(AV). The result can be treated as a word if and only if the AV score is occurred above threshold. Below Equation (1) use to calculate weighted score.

$$\text{Weight (xij)} = \frac{\log(ftij+1.0)\times\log(\frac{N}{PTi})}{\sqrt{\sum_j\left[\log(ftij+1.0)\times\log\left(\frac{N}{PTi}\right)\right]\times\left[\log(ftij+1.0)\times\log\left(\frac{N}{PTi}\right)\right]}} \text{ [19]} \quad (1)$$

For the making the transcription of the video, the full video is frame into passages and occurrence of number of times of passage has been calculated. Xij is the term i in passage j. frequency of i is represented by term ft. PT means the passage that having term i. The term in the denominator term is only normalizing factor.

The keyword tags highlight the content of searching repository [13,14], it helps the user to search the desire appropriate video. The benefits of using KVSUM and FF can be seen in operating user interface of this engine. The proposed KVSUM model generate the keyword as the below criteria.

$$F(xi) = fp_{min} + \frac{W(xi)-\min(W(xi))]\times[fpmax - fpmin]}{[\max(W(xi)) - \min(W(xi))]} \text{ [19]} \quad (2)$$

$$W(xi) = \sum_j Weight(xij)$$

F(xi) : font size of keyword xi

W(xi) : Weight(xij)

Earlier user is dependent on watching full video whether, however this videos are summarize to 10% of original length. This problem is improved by making user interface that allow to navigate the video at any position of timeline. This provide faster access to play video. The automation process of this model provide facility to browse the content easily. KVSUM come up with the feature like video cover, key frames, content summarization, thumbnails and keyword cloud. Video summarization module not only produces two or three result for thumbnail but also key from frames. User interface window also enables to play video on clicking the thumbnails.

The framework retrieves of visual information on the basis of feature extraction improves the effective search on video by extracting the feature of video on the basis of frame and image processing work on the video. Enhance the relevancy of video search with the query processing on the basis of metadata about the multimedia [21].

The keyword from the video on the basis or speech recognition engine help to understand the working principle of modern video search engine. Focuses on the enabling the jump on the scene of video by querying the keyword to help user to check relevancy of the video and to save user time [22].

The extraction of keyword which involves image processing techniques to summarize the video content. While other feature of media content like thumbnail, subtitle generation, annotation helps the learner to know about video. As user want more relevant video and hence this feature helps the engine to achieve more relevancy in its search operation [19]. Closed caption and Annotation makes the video more searchable was introduce by Microsoft's project Tuva [23].

The recommendation engine model has been improved engagement of user by recommending them related video that will increase in profit share with the help of advertising campaign. The proposed methodology returns a set of relevant videos by taking input which has metadata about user. System use the result *cookie* and list of viewed videos [24].

Chapter 3

Problem Statement

3.1 Problem Background

In the present scenario, learning through video content is most common. As the awareness of learning through video material is increasing, the more users are accessing video from VSE. Hence video views are increasing and growing at a huge level. Some websites rely more on video content. For instance NFL.com, site is loaded with lots of videos site is dominated by video content. Enterprises upload relevant videos on such promotional sites so that people can extract the most relevant video content. Though the VSE has a capability to match the query with the extracted keyword from the video, still VSE places the published video at the high rank on the higher number of result page.

3.2 Problem Statement

User normally choose top ranked and most views videos irrespective of usefulness. In order to improve video ranking, various video attributes have to be analyze. To increase more number of views, ranking is important factor. A simple and straight forward methodology is required to get the higher video rank by exploiting video search engines strategy.

During the literature survey, it was analyzed that there was lack of work done for suggesting such methodology. Eventually higher ranked video may have more views and as a result apparently promotes website visits. To promote white hat SEO, technique is suggested in this thesis to achieve the above goal by usage of keyword tags in title, description and transcript etc.

Chapter 4

Methodology

The step-by-step strategy to be followed in performing video search engine optimization. The following methodology reported improvement in ranking of video of search engine.

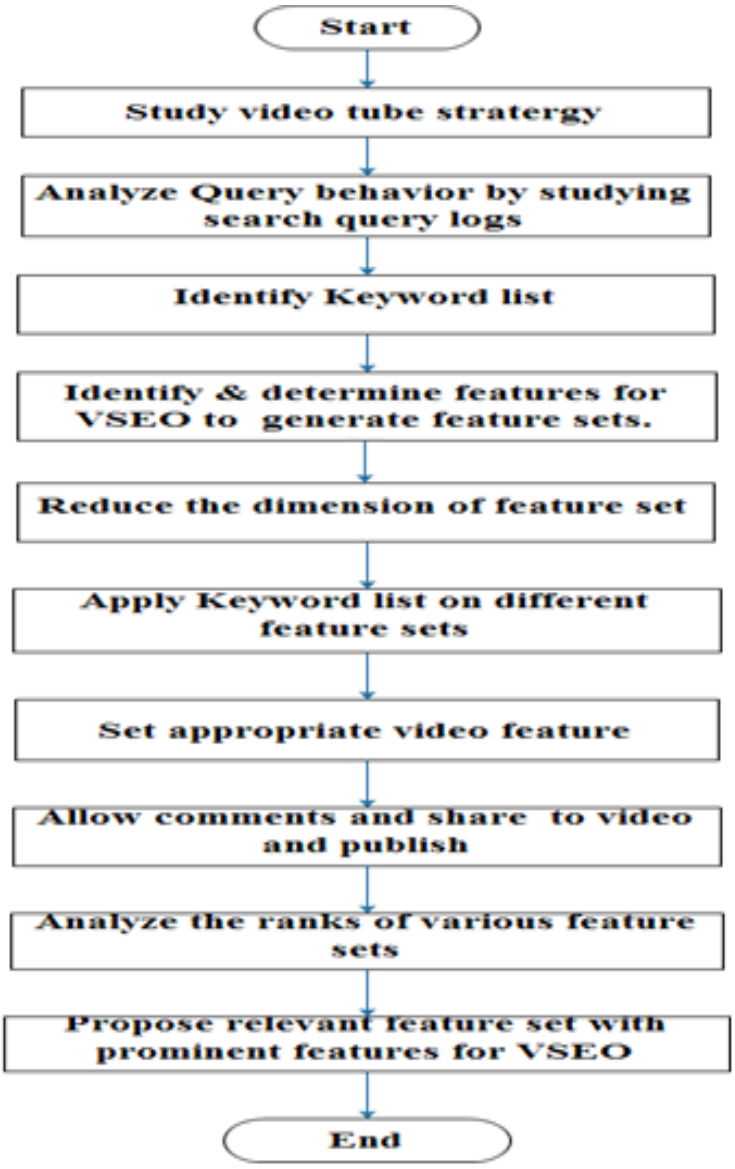


Figure 4.1: Methodology for VSEO

Steps followed for VSEO.

Step 1: Identify the basic search principle of video tube by searching with desire query and identify on which strategy video tube produces result.

Step 2: Analyze the type of video being published and accordingly identify query group on the basis of video category. Identify initial target query by studying query logs.

Step 3: Using video category, content and initial target query generated from Step 2 create the initial keyword tag lists.

Step 4: Identify and determine features from chosen VSE and generate feature set.

Step 5: Reduce the dimension of feature set by removing irrelevant features of set.

Step 6: Apply Keyword tag list on different feature sets in a way to avoid repetition and penalization.

Step 7: Generate and associate appropriate Thumb nail based on importance of video and business with thumbnail tag.

Step 8: Allow viewer to comment and promote video link to other social networking sites.

Step 9: Analyze the ranks with respect to different feature set. If search rank is not up to mark go to Step 2 else go to Step 10.

Step 10: Propose relevant feature set with refined features for VSEO.

Chapter 5
Implementation and Result

As per the proposed methodology, implementation started by choosing available different video search engine on the internet. There are different search engines like Youtube, Blinkx, Dailymotion etc. Most of the other available search engine returns result from youtube.com because Youtube has largest video data repository. For the study, youtube and Dailymotion has been taken as case.

5.1 Sample Cases Before Optimization

While studying the problem, some videos were uploaded. It was found that though the content of the video uploaded are correct and appropriate as per the query. Still videos have been placed at lower order of the result.

Case 1: "*best food spot in india*".

Figure shows case 1 video got placed in the result even query matches exactly with the title. Due to the huge amount of submissions of video on VSE, while searching video, it was found that among 3,920,000 videos case 1 is place at page 11.

Case 2: "*video lecture for the beginner*".

Video case 2 was uploaded on popular web video repository sites like Youtube, Dailymotion, etc. are often based on titles and description that publisher assign to their video. Figure shows that where video got placed in the result even if the query matches exactly with the title of the video. Due to the huge amount of submissions of video on VSE, while searching it was found that among 748,000 videos Case 2 video is placed at page 17.

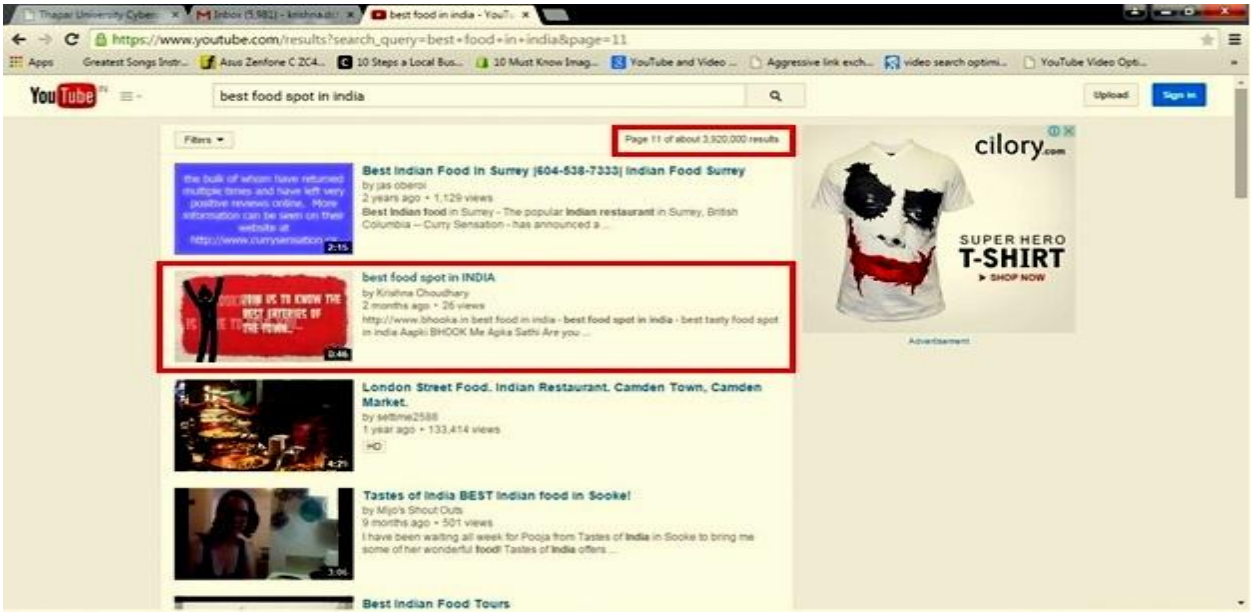


Figure 5.1: Result of search query “best food spot in india”

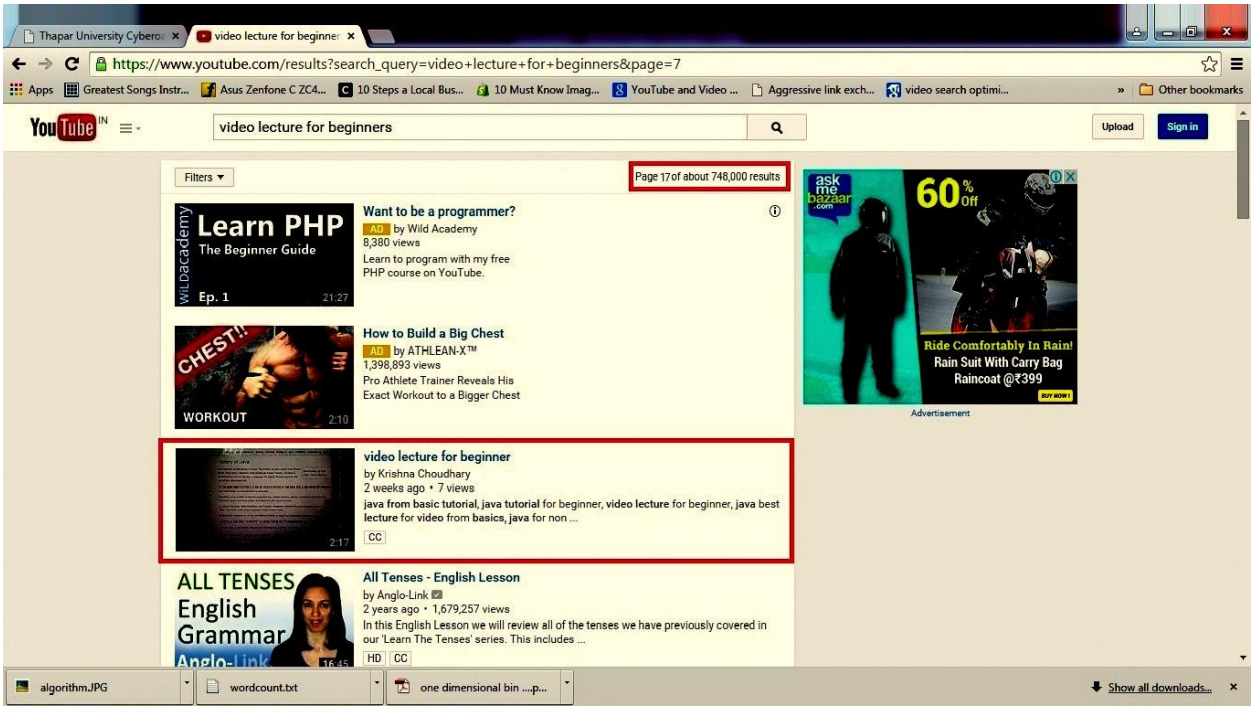


Figure 5.2: Result of search query “video lecture for beginner”.

Case 3: “best shots by Sachin Tendulkar”

This video was downloaded from the video tube and again uploaded to the same video tube with same title. The keyword chosen for the description was also same and also they

were heavily loaded with keyword that are related to content of the video. After searching with the same title and it has been placed somewhere in the result of about 27,200.

5.2 Study on Video Tube Search Strategy

On selecting the target video tube, study start with searching video search engine's searching criteria.

Following criteria is collected from Youtube video search engine.

Criteria for metric development.

- Keyword from the query found in Title of the Video (Yes/ No)
- Keyword from the query found in Description of the Video (Yes/ No)
- Total number of Views to the Video
- Video Age (Months Since Upload)
- Total Ratings collected from Both Positive & Negative response (i.e. like and dislike)
- Positive response rating in terms of Percentage give to the Video (Positive Ratings / Total Ratings)
- Total Comments
- Video length(video Time)

Query: Best food spot India.

Table 5.1: Analysis of Youtube search result.

Query	Title	Description	Views	Age in months	Rating	Rating %	Comment	length
Result 1	Yes	Yes	124,536	4	211	82.93	20	12:30
Result 2	NO	NO	1,94,970	24	2284	84.10	900	4:29
Result 3	Yes	NO	360,243	36	381	67	181	5:23
Result 4	NO	YES	24,486	10	36	83	7	1:36
Result 5	Yes	YES	90,399	24	1261	99.36	155	4:19

Table 5.1 shows that result generated on searching of query "best food spot in India". This data has been collected from the first page of the result.

On the basis of Table 5.1 the following observations have been made before optimization.

1) Youtube result are not heavily based on title and description because Result 2 having no keyword in title and description.
2) From the Result 2 it can be inferred that age of the video can be the factor for the ranking in the place second of the first page but the Result 1 shows that video was uploaded before 4 month.
3) Number of views can also be the important factor for showing the result at very first page but Result 4 has less view in terms of number.
4) Positive rating percentage is also the important factor for gaining the trust about the Video but Result 3 doesn't clear the idea of placing it on number 3 as the positive rating of the video is poor as compare to other.

5.3 Strategy for Selecting Target Audience and Keyword

According to the type of video, identified the targeted audience by using Google adword tool's Display planner.

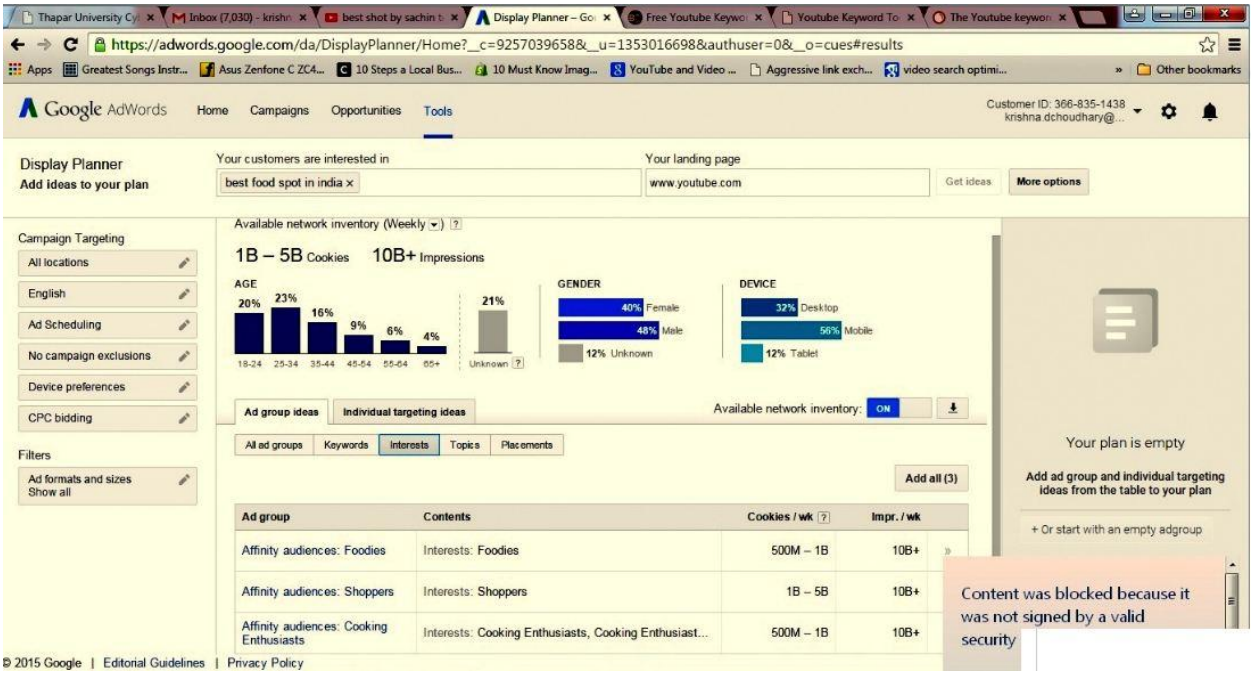


Figure 5.3: Keyword research thought Google Display planner.

Above Figure 5.3 shows that Google display planner can be used to get ideas for keywords, placements, and all other Display Network targeting methods help for ad campaign.

It help to take decisions and show estimation of past result of how many types of user devices used those keyword.

5.4 Psychology Analysis of User Search Queries

In the searching activity, SEO professional concentration on keywords but understand the behavior of searching query by the user also necessary. The basic three category in which searches are divided.

5.4.1 Navigations Search Queries

In this type user enters other website name into search engine rather than directly feed into URL. Example: user enters “olx” into search engine rather that entering into URL, www.olx.com. Google categories this searches as “go queries”. This is not always the case, user might be entering “olx” in search not to visit the website but only to read its news and information [25].

5.4.2 Informational Search Queries

Generally user enters long descriptive query to get any particular target website which called as informational queries. Mostly they are looking for some information or probably they need the answer to questions. Google makes knowledge graph to show the result for such queries [27].

5.4.3 Transactional Search Queries

This type of query include transactional keywords or user enters some queries to perform transaction after purchase of the product. For example user search for “Lenovo laptop” and also used “buy”, ”purchase”, “order”. Study found that most of the times, people click on the paid sites rather than unpaid result [26].

Strategy to identify and determine keyword for applying to features of the video:

- As per the content of the video, identify the key phrases and sentences, lines, important word about the video and set the target title or main search query.
- Remove the stop words, apply the word stemming on identified lines for the video obtain keyword list.
- Check the term relevancy according to the target search query.
- Create the relevancy list related to the line by line percentage of relevancy.

First, there is the video's relevance. This is a measurement of how relevant the video is to the search query. In YouTube, since the content is video, the best measurement of relevance is the video title.

A good way to measure the relevance is to count the number of words in the actual video title that match the actual query. For example, say the target search query is "Best food spot in India".

Some example keyword list for the video can be rated as follows:

1) best tasty food spot in India = three words matches from the whole query. So it is 4/5 = 80% relevant. The query contains four words.

2) best hotel food item list in India = Only three word matches from the whole query. So it is 3/6= 50% relevant

3) tasty food spot at Patiala = 2 words match the search query, which means it is 2/4 or 80% relevant

Table 5.2: List of keyword list and relevancy

Keyword list	Relevancy in %
1	100
2	80
3	80
4	50
5	50

- Prepare to place these keyword list in different set of feature set of the video search engines site.

5.5 Identification and Determination of Different Feature Set

On the basis of video search engine's feature, created a feature set. Basic main feature set consist of all the features {Title, Description, Tags, Thumbnail, Transcription, Annotation, Share, and Comment}. Feature vector (Fv) = {f_1, f_2, f_3, f_4, .. ,f_n}

Fv = $\bigcup_{i=1}^{n} fi$ & fi $\in Fv$

F_1= {Title, Description},

F_2= {Title, Description, Tags},

F_3= {Title, Description, Transcription},

F_4= {Title, Description, Transcription, Annotation},

F_5= {Title, Description, Transcription, Annotation, Comment, Share}.

Keyword set are derived from keyword suggestions by Google, Youtube and keyword suggestion tool.

Keyword set (Ks) = {Kl_1, kl_2, kl_3, . . , kl_n}

Ks = $\bigcup_{i=1}^{n}$ kli & kli∈ Ks

The experiment was performed on the basis of 7 video set and collected the result. Among those experiment, Case 2 which is uploaded with

Title “video lecture for beginner”.

kl_1 = {video, lecture, beginner},

kl_2 = {java, tutorial, video, lecture},

kl_3 = {best, java, video, lecture, beginner, certified, trainer, non CSIT},

kl_4 = {java, lecture, beginner, programming, language}.

Ql is the query list that based on psychology of analysis of user search queries.

Ql = $\bigcup_{i=1}^{n} Qli$

Ql = {Ql1, Ql2, Ql3,.. , Qln}

FS – kli = optimize feature set (Fs) with keyword list (kli).

5.6 Optimization Result of Youtube Video

Study shows that keyword selection are important. This is the case in which video name "video lecture for beginner" has been uploaded to www.Youtube.com.

Table 5.3 shows the placement of the keyword in the feature improves the ranking of the video significantly. Placement of most appropriate keyword k3 inside the feature set f5 shows the change in the rank of video significantly as compare to the other feature set.

Table 5.3: Keylist vs Features from Youtube

kli\fi	F1	F2	F3	F4	F5
kl1	416	411	382	382	**351**
kl2	260	251	238	238	**211**
kl3	84	64	34	34	**16**
kl4	619	612	579	579	**560**

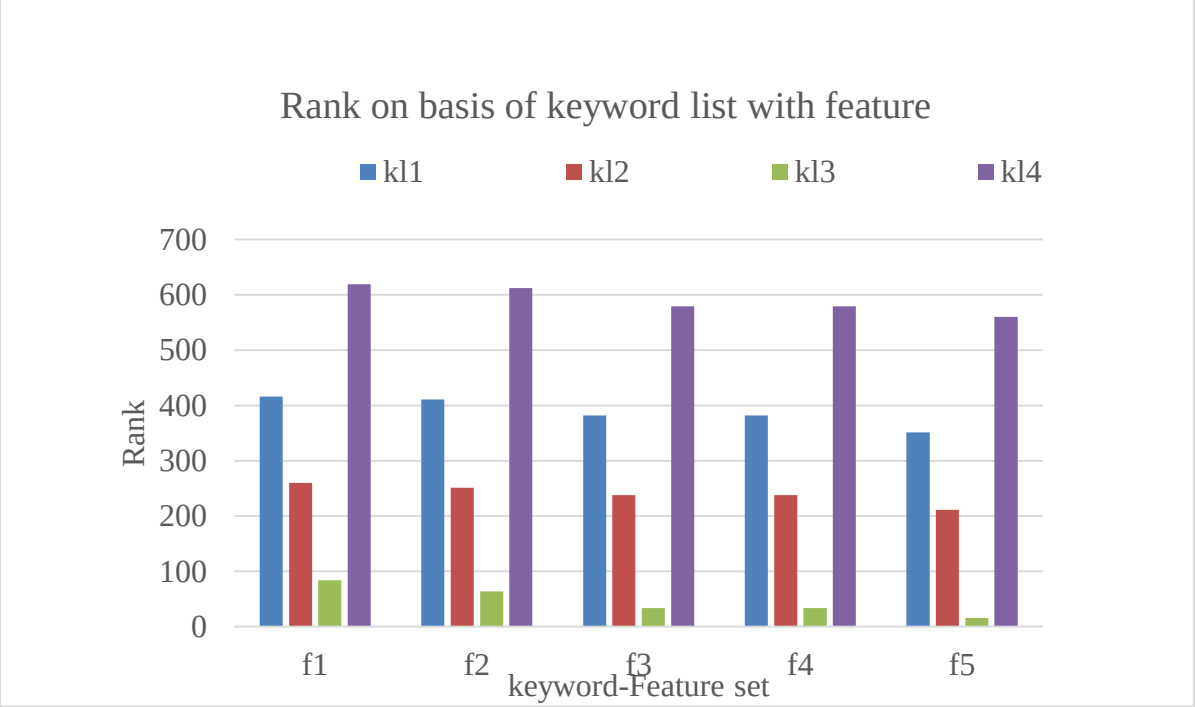


Figure 5.4: Analyze of rank on the basis of keyword list vs feature.

In Figure 5.4 shows that from keyword list kl1 with feature f1 to kl5 with f5, in which keyword list kl3 with feature f5 is results with most significant improvement in the rank. Decreasing status of the bar graph represents improvement in the rank.

Table 5.4 show ranking of video with different queries. Video attribute like Share and comments helps the search engine to gain the trust for the video. Table gives performance of combination of fully feature set against different Query.

Table 5.4: Query vs Optimize feature set– keyword list from Youtube

Qi\Fs – kl	Fs - kl1	Fs – kl2	Fs – kl3	Fs – kl4
Q 1	215	245	**7**	345
Q 2	226	226	**35**	360
Q 3	351	211	**16**	560
Q 4	254	207	**13**	553
Q 5	401	364	**159**	721

In Figure 5.5 the gray color line shows the combination of fully optimized feature set and keyword list k13 which says that once the combination is established in ranking then video search engine place that video on the top of pages even if query changes relevantly.

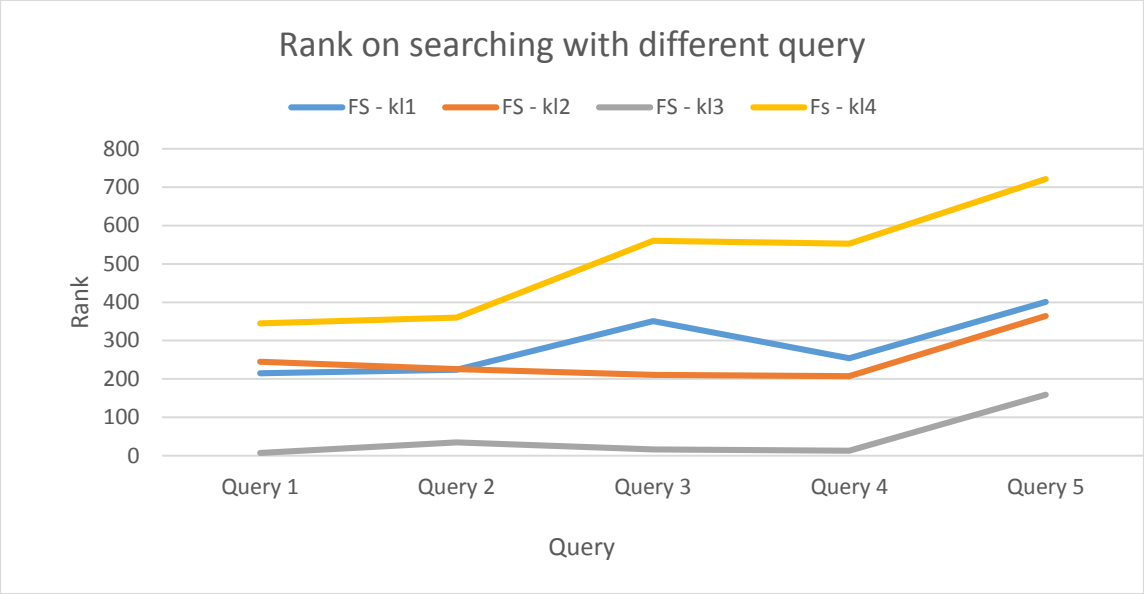


Figure 5.5: Analysis of the different query with optimize feature set & keyword relation

5.7 Optimization Result of Dailymotion Video

Another popular video search Engine which has more number of submission of video is Dailymotion.com. The same video with same title name had uploaded on the Youtube.

Feature sets for Dailymotion site.

F_1= {Title},

F_2= {Title, Description},

F_3= {Title, Description, tags}

The keyword list same as the video uploaded on the Youtube.

Table 5.5 shows the Placement of the keyword in the feature improves the ranking of the video significantly. Though the result of Dailymotion.com is heavily based on keyword in the title.

Table 5.5: Keylist vs Features from Dailymotion.

Kli\fi	F1	F2	F3
K1	4	4	4
K2	94	90	90
K3	**1**	**1**	**1**
K4	42	42	42

In Figure 5.6 shows that Decreasing status of the bar graph represents improvement in the rank.

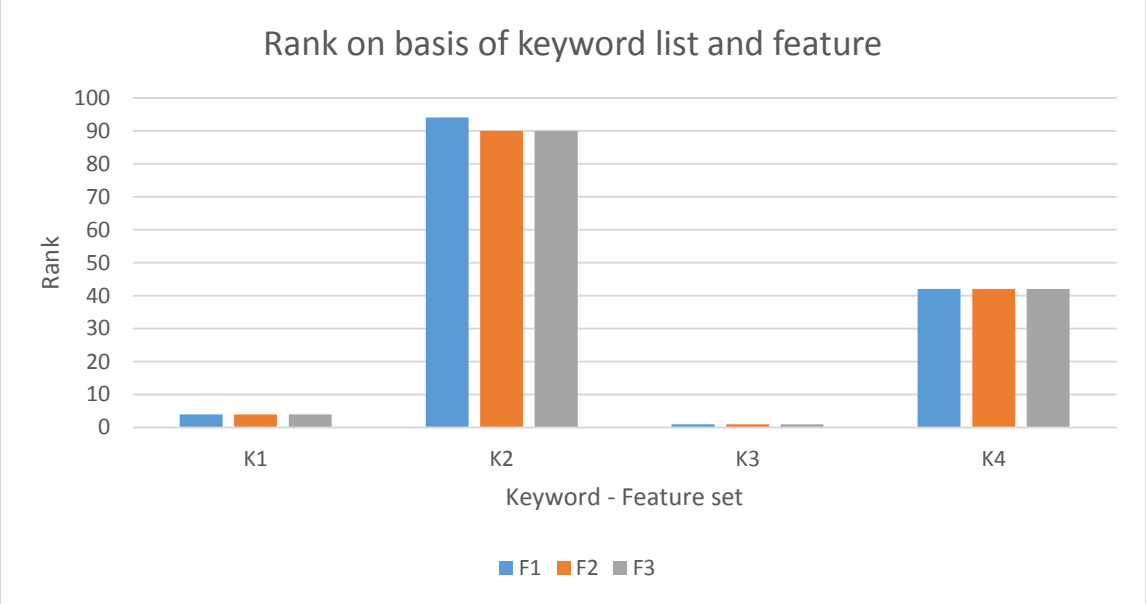


Figure 5.6: Analysis of rank on the basis of keyword list vs feature.

Table 5.6 show ranking of video with different queries. Those result are in top which having high relevancy of title with the search query. As the relevant keyword from the title matches with the search query will move the video to the top searches.

Table 5.6: Query vs Optimize feature set– keyword list from Dailymotion

Qi\Fs – kl	Fs - kl1	Fs – kl2	Fs - kl3	Fs – kl4
Q1	42	47	**1**	49
Q2	79	79	**83**	85
Q3	**4**	**90**	**1**	**42**
Q4	80	92	**62**	50
Q5	101	110	**56**	107

In Figure 5.7 the lowest line shows the combination of fully optimized feature set and keyword list k13 which again moving video to the top search. Bottom most gray line indicating that improvement in the ranking can be gain through optimize feature set and keyword list.

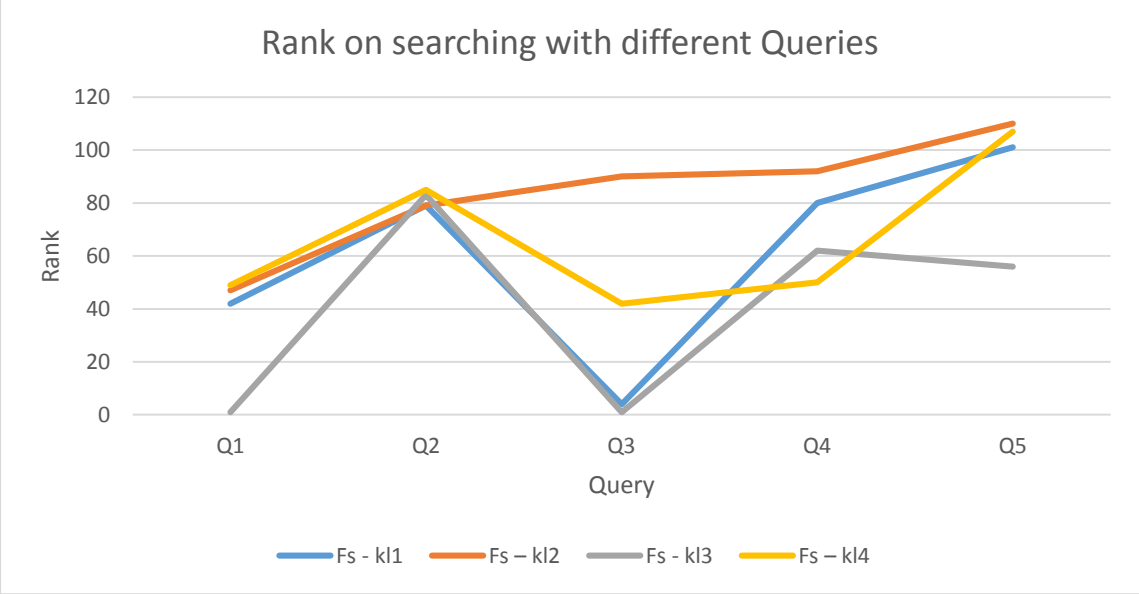


Figure 5.7: Analysis of the different query with optimize feature set & keyword relation.

Table 5.7 shows the effect of placing the unique keyword inside the description of the video. In this experiment, each line of the Description is considered as query. With the query 4 video has been placed at top because query 4 has unique keyword.

Table 5.7: Rank with the unique keyword in the description.

Query as keyword list	Rank with uniqueness in the keyword list	Unique keyword
Query 1	7	NO
Query 2	20	NO
Query 3	42	NO
Query 4	1	YES

Table 5.8 shows video search engine gains the trust rank through matching the content of description with comments. Comments are the important factor to gain trust because viewer comment on the video and which give idea to the video search engine that content are same as mentioned by the publisher. Video search engine validate the content of video with comments posted by viewer.

Table 5.8: Description and comment matching criteria for ranking of video.

Best food spot India	Title	Description	Description = comments
Rank 1	Yes	Yes	Matching
Rank 10	NO	NO	Matching
Rank 40	Yes	NO	Matching
Rank 90	**NO**	**YES**	**Not matching**

5.8 Result After Optimization

In one of the case, mentioned in figure 5.2, the improvement in rank is shown in Figure 5.8. Video is placed in page 1 of results and is sandwiched between videos which are older

and have more number of views. Following result shows that placing of keyword inside the feature in of video improves the ranking.

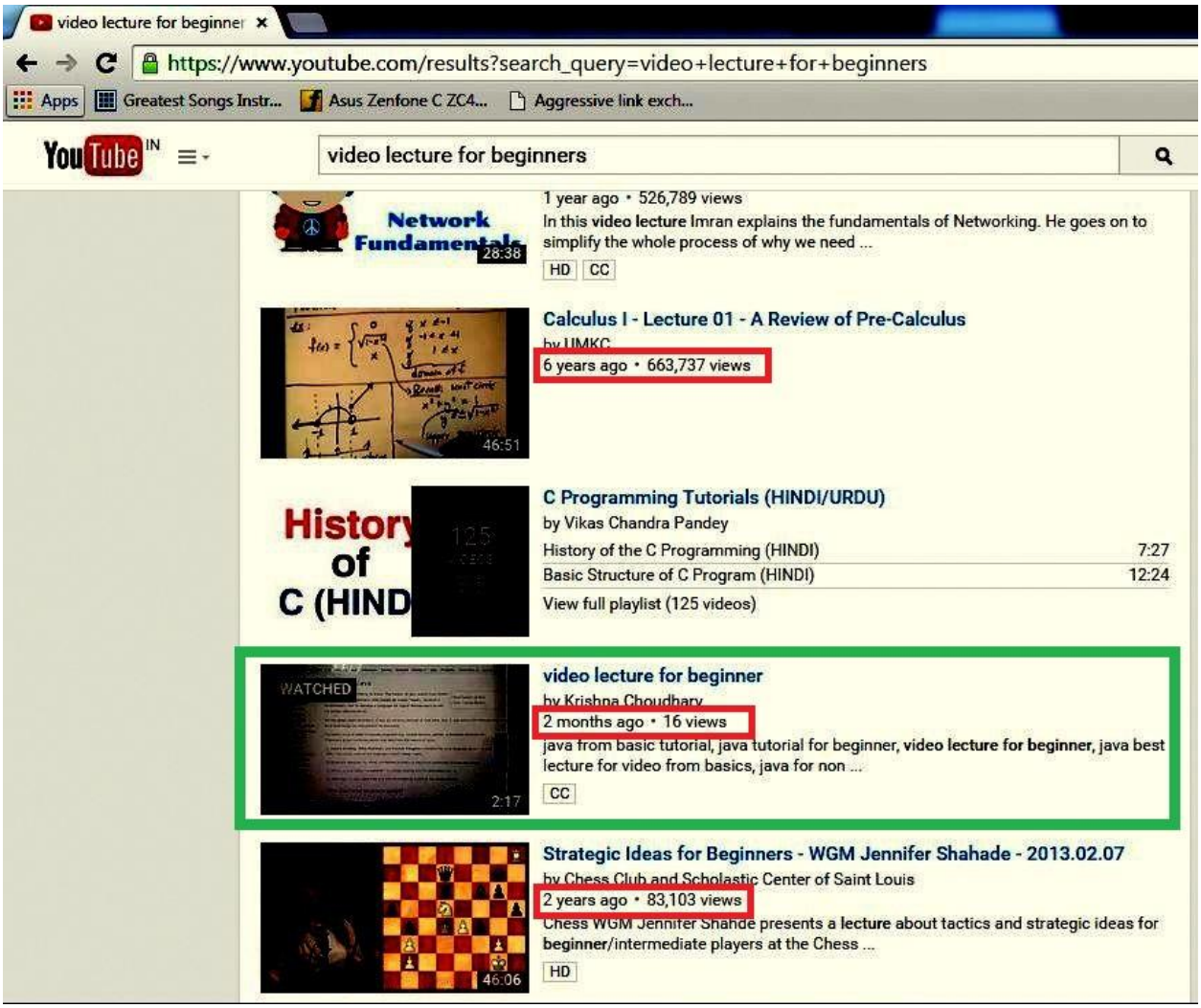


Figure 5.8: Result after performing VSEO on video.

In one the case the video was uploaded with search keyword “best shots by Sachin Tendulkar” is found panelized by the search engine. As the video content was not original, it contains delicacy also it was downloaded from Youtube.

Ignoring the rules proposed by the search engine may lead to penalties such as reducing the search ranking and in some exceptional cases completely banning the video by the search engine.

5.9 Suggestions for Optimization

On the basis of optimization result the following observations have been made after optimization.

1) Figure 5.3, 5.4, 5.5 and 5.6 shows that placement of keyword inside the feature are most important than any other factor.
2) Figure 5.7 demonstrate the fact that age of video is not as much deciding factor for getting into top pages.
3) Number of views can also be the important factor for showing the result at very first page.
4) Table 5.8 shows the VSE uses comments and description to gain trust rank of video.
5) Table 5.7 shows the usage of unique keyword can place the video on top place.
6) On observing Case 3 results an important conclusion is drawn that do not duplicate or copy content of other videos otherwise search engine may penalize and drastically reduce the ranking of video in spite of applying optimization technique.

5.10 Suggested Prominent Feature

On the basis of proposed research, prominent features were identified for the Video Search Engine Optimization.

On site optimization

- Keyword research

  Keyword analysis is the most important part of optimization. As in SEO of website do keyword research for online marketing, for that one has to target audience of video and have to do research like what audience will place the keyword inside the search engine. This keyword analysis is provided by some keyword analysis tool like Google Adwords, Youtube keyword suggestions or query log analysis.

- Video tags

  While uploading any video to any video search engine, keep that video name same

relevant to the keyword. Keeping the different video name irrelevant to the keyword like vo231232.mov will create ambiguity to the robot of video search engine.

- Title

  Title keeps important meaning to place a video in the top search because most the search engine searches the queries related to titles. Place the keyword in the title will help the viewer to click on the video as it seems to be most related video according to their search. Keep the title that catch the eye of viewer to click it on.

- Description

  Describe the video is all about to create the interest to viewer to watch the video. Viewer gets the idea related to details of the video and it is deciding factor for viewer to view that video or not. Describe about company location of the shooting of video also put information related to the company like location, Date of event Name, phone number. First line of description should be companies URL will help to direct redirect on the company's website. The information inside the description should not be so large to read as viewer came to watch video not the whole description about the video.

- Thumbnail

  Frames from the videos are treated as thumbnail. This are the auto generated thumbnail but mostly chosen thumbnail are preferred. It the first look of the video and gives the picture about what does inside the video. This will be key factor to inviting the viewer.

- Video Transcript

  Many video search engine gives the facility to put transcript the video. Many engines check the transcript to help to suggest the related result to the user. It is just like timeline or subtitle of the video so also should put the keyword inside the transcript.

- Annotations

  Annotations provide facility of embedding external links to publisher's website. It is an important aspect, as it links to website and will either be shown during full video length or as per the time of appearance set by the uploader. Thus if viewer likes video and decides to go to the publisher website then providing this direct link on the video will increase the visits to the website.

Off site optimization

- Views

  This is main motive of video search engine optimization. After considering rank of the video in the search result, next factor is views that the video has. Larger the views, larger is the chance of viewer to watch the video and this invites more traffic.

- Shares

  The presence of video on the social networking sites like Facebook, twitter, Google+ will invite more number of views to the video. Many people don't go to search engines to watch the videos, they might like to watch those video that are shared by friends, present on some social website.

- Comments

  This guide the search engine that whether the video got attention of viewer till the end. It help the engine to know that video having appropriate content.

- Channel Authority

  If all the videos are link with Publisher's channel then people are more likely to subscribe the channel. This also help search engine to recognize that videos from the channel are more authenticated.

Chapter 6
Conclusion

In this study, Analysis has been done on video tube feature for exploring the area of video search engine optimization (VSEO). Study shows the comparative analysis of before and after VSEO. Main focus was on video search optimization, proposed methodology helps to identify key attributes for a video while searching. VSEO is keyword selection strategy, which is deciding factor for the ranking of video. Various feature sets were identified during research work on which keywords were applied to improve the ranking. Experimental results shown significant improvement in the ranking as per methodology. This technique has been applied on popular video search engine website like Youtube and Dailymotion.

Usually search engine uses video feature detection strategy on the video to get its basic information of content of the video. Placement of appropriate keyword are important to boost the ranking. Understanding of user query helps to identify keyword that are essential for video. It has been observed in the research work that due to over usage of the keyword by publisher, the video can be penalized by search engine. It can also happen when the content is not original and trust worthy.

Chapter 7

Future Work

Ever changing technology, tools, exponential increase in the number of user and number of videos on video tubes influences video search engines to keep on changing their search ranking strategy.

Video Search Engine Optimization is ongoing process. There is no proper optimization process for improvement in the rank which is sustained for longer time. Video tube might add additional properties and parameter for uploaded video. Those properties can be accommodated in feature set to improve the ranking. Moreover image processing techniques can be applied on thumbnails and video frames for improving the search ranking.

Chapter 7

References

[1] P. Arora and T. Bhalla , "*A Synonym Based Approach of Data Mining in Search Engine Optimization*", International Journal of Computer Trends and Technology (IJCTT), Vol. 12, pp. 201-205, 2014.

[2] Killoran and J.B, "*How to Use Search Engine Optimization Techniques to Increase Website Visibility*", The IEEE Transactions on Professional Communication, Vol. 56, pp. 56-66, 2013.

[3] S. Zhang and N. Cabage, "*Does SEO Matter? Increasing Classroom Blog Visibility through SearchEngine Optimization*", In Proceeding of 46th Huawaii International Conference on System Sciences (HICSS), pp. 1610 – 1619, 2013.

[4] R. Malanga, "*The Value of Search Engine Optimization: An Action research Project at a New E-commerce site*", Journal of Electronic Commerce in Organizations, Vol. 5, 2007.

[5] D. Dover and E. Dafforn, "*Search Engine Optimization (SEO) Secrets*", pp. 55-70 2011.

[6] Invoedo, E-Commerce Video Benchmarks Report: Q2 & Q3 2013 [Online]. Available: http://www.invodo.com/wp-content/uploads/2013/12/Invodo-Video-Benchmark-Report-2013-Q2-Q3.pdf, 2013.

[7] eMarketer, As Barriers Tumble, Video Marketing Adoption Grows [Online]. Available: http://www.emarketer.com/Article/Barriers-Tumble-Video-Marketing-Adoption-Grows/1010374, 2013.

[8] eMarketer, Do Millennials Watch Online Videos Differently? [Online]. Available: http://www.emarketer.com/Article/Do-Millennials-Watch-Online-Videos-Differently/1010444, 2013.

[9] The Most (and Least) Effective Keywords in Email Subject Lines [Online]. Available: http://www.marketingprofs.com/charts/2013/11228/the-most-effective-keywords-in-email-subject-lines, 2013.

[10] Google Search, Importance of Video In Online Marketing [Online]. Available: http://www.reelseo.com/video-email-effective-survey/#download, 2013.

[11] T. Tuna, J. Subhlok and S. Shah, "*Indexing and Keyword Search to Ease Navigation in Lecture Videos*", In Proceeding of Applied Imagery Pattern Recognition Workshop (AIPR)**,** Vol. 1,pp. 1-8, 2011.

[12] C.G.M. Snoek, K.E.A. van de Sande, O. de Rooij, B,Huurnink, J.C. van Gernert, J.R.R.Uijlings, J. He, X. Li, I. Everts, V. Nedovic, M. van Liempt, R. van Balen, F. Yan, M.A. Tahir, K. Mikolajczyk , J. Kittler, M. de Rijke, J.M. Geusebroek, Th. Gevers, M. Worring, A.W.M. Smeulders and D.C. Koelma, "*The MediaMiII TRECVID 2008 Semantic Video Search Engine*", ISLA, University of Amsterdam Amsterdam, the Netherlands, 2008.

[13] A. Amir, J. Argillander, M. Campbell, A. Haubold, G. Iyengar, S. Ebadollahi, F. Kang, M. Naphade, A. Natsev, J. Smith, J. Tesic and T. Volkmer "*IBM Research TRECV/D- 2005 Video Retrieval System*", TREC Video Retrieval Evaluation Online Proc, 2005.

[14] S. Chang, W. Hsu, W. Jiang, L. Kennedy, D. Xu, A. Yanagawa, and E. Zavesky "*Video Search and High-Level Feature Extraction*", TREC Video Retrieval Evaluation Online Proc., 2005.

[15] Y. Song and G. Marchionini, "*Effects of Audio and Visual Surrogates for Making Sense of Digital Vide*o", In Proceeding of the SIGCHI conference, pp. 867-876, 2007.

[16] G. Marchionini, Y. Song, and R. Farrell, "*Multimedia surrogates for video gisting: Toward combining spoken words and imagery*", Inf. Process. Manage, Vol. 45(6), pp. 615-630, 2009.

[17] M. B. Wildemuth, G. Marchionini, T.Wilkens, and et.al, "*Alternative Surrogates for Video Objects in a Digital Library: Users' Perspectives on Their Relative Usability*", In Proceeding of 6th European Conference on ECDL, Vol. 2458, pp. 493-507, 2002.

[18] B. Wildemuth, G. Marchionini, M. Yang, G. Geisler, T. Wilkens, A. Hughes, and R. Gruss, "*How fast is too fast?: evaluating fast forward surrogates for digital*

video", In the Proceeding of the 3rd ACM/IEEE-CS joint conference on Digital libraries, pp. 221-230, 2003.

[19] W. Chang, J. Yang, Y. Wu, "*A Keyword-based Video Summarization Learning Platform with Multimodal Surrogates*", In the Proceeding of Applied Imagery Pattern Recognition Workshop (AIPR), pp.1-8, 2011.

[20] A. Yousef, Md. Chatti and U. Schroede, "*Video-Based Learning: A Critical Analysis of The Research Published in 2003-2013 and Future Visions*", In Proceeding of 6th International Conference on Mobile, Hybrid, and On-line Learning, Vol. 4, pp. 112-120, 2014.

[21] P. Geetha and V. Narayanan, "*An Effective Video Search Re-Ranking for Content Based Video Retrieval*", In Proceeding of 3rd International Conference on Trendz in Information Sciences and Computing (TISC), pp. 55-60, 2011.

[22] T. Yoshida, K. Tada and S. Hagai, "*A Keyword Accessible Lecture Video Player and Its Evaluation*", In Proceedings of International Conference on Information Technology: Research and Education (ITRE2003), pp. 610-614, 2003.

[23] Microsoft Project Tuva [Online], http://research.microsoft.com/apps/tools/tuva/.

[24] A. Chatterjee, B. Sarkar, P. Chandraghatgi, K. Seal, G. Ananthakrishnan, "*Search Based Video Recommendations*", In proceeding of 4th National Conference of IEEE on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Vol. 10, pp. 1-5, 2013.

[25] Mothe, B. Rohini, Deshmukh and V. S, "*A novel method to Automatically Categorizing Search Results using Web Search Goals*"**,** International Journal of Computer Applications, Vol. 98, pp. 22, 2014.

[26] I. Kang, "*Transactional Query Identification in Web Search",* In proceeding of Springer on Second Asia Information Retrieval Symposium, AIRS 2005, Vol. 3689, pp. 221-232, 2005.

[27] B. Jansen, D. Booth and A. Spink, "*Determining the informational, navigational, and transactional intent of Web queries*", International Journal of Information Process and Management, Vol. 44, pp. 1251-1266, 2008.

Chapter 9

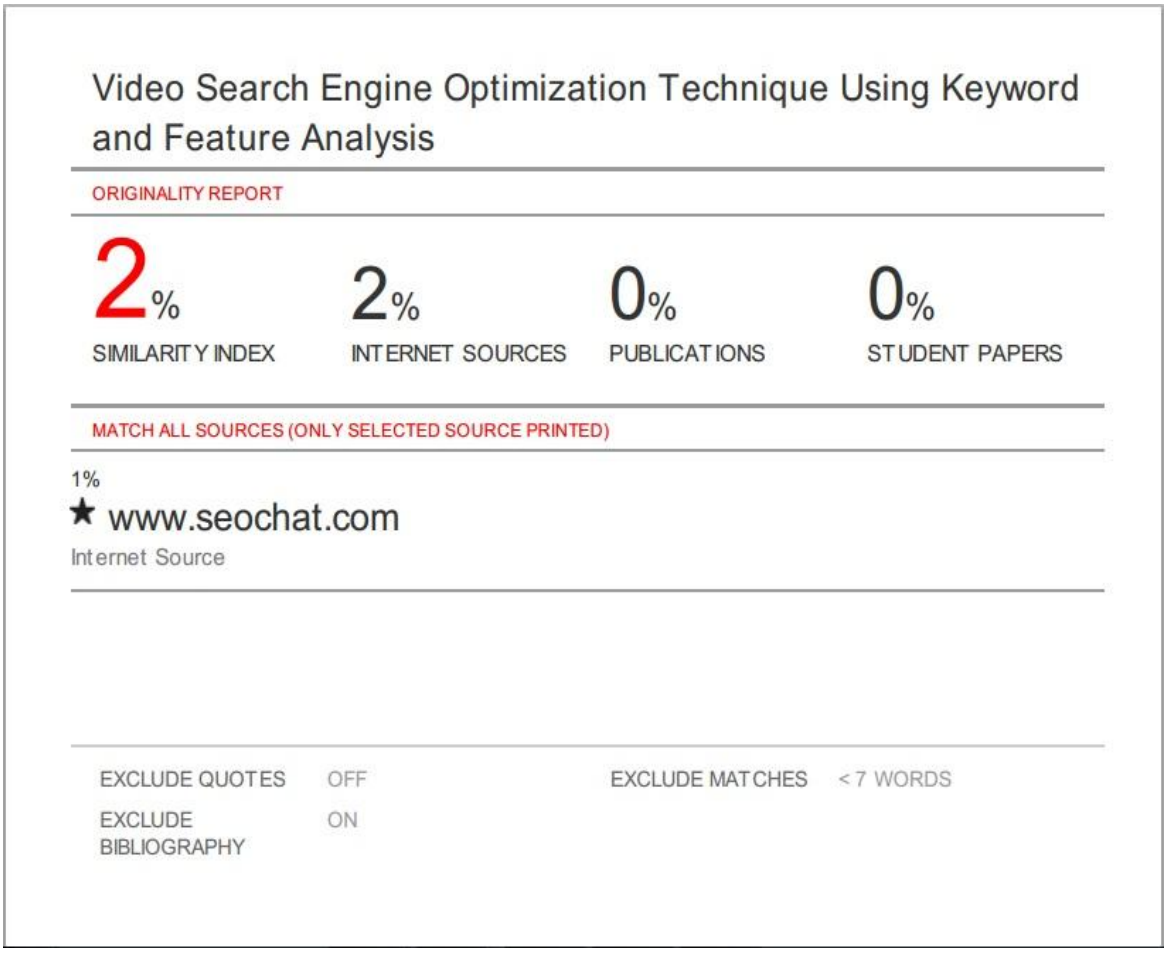
List of Publications

K. Choudhari, V. Bhalla, "Video Search Engine Optimization Technique Using Keyword and Feature Analysis", In Proceeding of Second International Symposium on Computer Vision and the Internet (VisionNet'15), Kochi, Kerala, 2015 [Accepted].

Youtube link : http://youtu.be/WcJ2pUS30u8

Plagiarism Report

Following is the snapshot of plagiasm report of above work.



Video Search Engine Optimization Technique Using Keyword and Feature Analysis

ORIGINALITY REPORT

2%	2%	0%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

1%

★ www.seochat.com

Internet Source

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	< 7 WORDS
EXCLUDE BIBLIOGRAPHY	ON		