TECHNOLOGIES FOR CONTEXT BASED VIDEO SEARCH

A Thesis
Presented to
The Academic Faculty

by

Arshdeep Bahga

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2010

TECHNOLOGIES FOR CONTEXT BASED VIDEO SEARCH

Approved by:

Professor Vijay K Madisetti, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Sudhakar Yalamanchili
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Shamkant B Navathe
College of Computing
Georgia Institute of Technology

Date Approved: 24 March 2010

To my parents and my brother

for their unconditional love and support.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my advisor Professor Vijay K Madisetti, for his support, patience and motivation during the last two years we have been working together. His guidance has helped me throughout my research and also in writing of this thesis.

Professor Sudhakar Yalamanchili and Professor Shamkant B Navathe deserve a special thanks for serving as my thesis committee members, and giving me valuable insights into my thesis topic.

I would also like to thank my colleagues Adeel Yucef and Simeranjit Brar, for their help, support and friendly discussions at work.

Finally, I would like to thank my parents and brother for their love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis presents methods and a system for video search over the internet or the intranet. The objective is to design a real time and automated video clustering and search system that provides users of the search engine the most relevant videos available that are responsive to a query at a particular moment in time, and supplementary information that may also be useful. The thesis highlights methods to mitigate the effect of the semantic gap faced by current content based video search approaches. A context-sensitive video ranking scheme is used, wherein the context is generated in an automated manner.

CHAPTER I

INTRODUCTION

Videos have become a regular part of our lives due the recent advances in video compression technologies, availability of affordable digital cameras, high-capacity digital storage media and systems, as well as growing accessibility to high speed communication networks and computers. Thousands of new videos are being uploaded over the Internet every second. However, without a fast and reliable video search engine it is difficult to retrieve videos.

The videos content available on the web ranges from news videos, video lectures on various subjects, music videos, etc. Some types of video content, particularly news video content is highly dynamic in nature, as different video news broadcasting website constantly keep uploading news videos. The current commercial web search engines are guided by textual content. These search engines crawl the web for new content, and index keywords from the text to make the content searchable. However, currently no commercially available search engine addresses the problem of searching dynamic video content such as news videos. The life time of news videos is only a few hours as new videos on the same news story are uploaded on the web every few hours. The video search engines available currently rely on the metadata information which is in textual form. This meta information is available in the form of video title, captions, descriptions and the textual content surrounding the video on a webpage. Although the metadata information is valuable for video search, however the amount of metadata available maybe limited and subjective in nature. Therefore, there is a need to extract more information related to the video. Current research efforts for video search have focused on augmenting the video meta-information with textual information from closed captions and automatic speech recognition. These approaches are error prone and therefore relying completely on this augmented meta-information may lead to incorrect search results.

1.1 Challenges in Video Search

Video search is much more complicated than text search, which has led to lot of research efforts in this fied. The challenges involed in video search are as follows:

1.1.1 Video acquisition

The way text search engines acquire new content is that they use crawlers to find content for indexing HTML pages. Links that are found in a crawled HTML page are used to crawl and index more pages. Acquiring video content on the other hand is more complicated as videos are not directly embedded in HTML pages. Most video broadcasting websites provide video content through streaming. Therefore direct links to a video may not be available by just parsing the HTML pages in which the videos are streamed. Videos are available is a number of formats such as Flash, MP4, etc. Therefore videos acquisition and indexing requires an additional step of video transcoding to convert videos in different formats to one standard format. Parsing videos to extract video metadata is another challenge. Unlike, HTML pages which can easily be parsed to extract the text from webpages, parsing videos is difficult, as the videos are available is a number of different formats, and have large number of parameters and supplementary information as a part of the file headers. The video metadata is often lost in transcoding, therefore the metadata extraction has to be done before just after video acquisition.

1.1.2 Video Indexing and Ranking

Ranking videos based on relevance to a query is more complicated than ranking text documents. A lot of algorithms are available for ranking and indexing text documents, and the relevance to a search query is done based on the number of matching keywords between the query and the text document. The video metadata information available is often incomplete and subjective. Therefore, keyword based ranking for videos may lead to incorrect results.

CHAPTER II

PREVIOUS WORK

There are many different approaches to video search, as discussed below:

2.1 Content Based Video Search

There are two categories of content based video search approaches which either use the low-level visual content or high-level semantic content, as described below:

2.1.1 Low-level visual content based search

The low-level content based approach uses low-level visual content characterized by visual features such as color, shapes, textures, motion, edges, etc for video search. These low level features can be extracted automatically to represent the video content. There are several different classifications schemes for video content. For example, MPEG-7 is a multimedia content description scheme, which has standardized more than 140 classification schemes that describe properties of multimedia content. MPEG-7 provides different descriptors for color, motion, shapes, textures, etc to store the features extracted from video in a fully standards-based searchable representation. Other multimedia description schemes used in the past are Thesaurus of Graphical Material (TGM-I), TV-Anytime, SMPTE Metadata Registry from Society of Motion Picture and Television Engineers and P/Meta Metadata Scheme from European Broadcasting Union. A limitation of low-level visual content based approach is the semantic gap between users queries and the low-level features that can be automatically extracted. Virage video engine [26], CueVideo [27] and VideoQ [28] are some of the low-level content based video search engines.

2.1.2 High-level semantic content-based search

The high-level semantic content based approach uses high-level semantic content characterized by high-level concepts like objects and events for video search. Unlike, low-level features which can be automatically extracted, the high-level semantic content is difficult to characterize from raw video data. The reason being that at physical level, a video is nothing but a temporal sequence of pixel regions without a direct relation to its semantic content. There are two different types of high-level semantic content based approaches:

2.2 Concept-based video search

The concept based video search approaches use concepts detectors (like building, car, etc) to extract semantics from low level features [16]-[20]. These use shared knowledge ontology such as WordNet or external information from Internet to bridge the semantic gap between the user queries and raw video data. For example, LSCOM (Large-Scale Concept Ontology for Multimedia) [21] includes 834 semantic concepts. MediaMill [23] extended the LSCOM-lite set by adding more high level semantic features (HLFs) amounting to a total of 101 features. Informedia [23] is another well known system which uses HLFs for video search. Though semantic concepts are useful in retrieving shots which cannot be retrieved by textual features alone, the search accuracy is low. To overcome these limitations, event based approaches have been used.

2.3 Event/Topic-based video search

The event/topic based approaches use event/topic structures from video for providing additional partial semantics for search. Text annotations, closed captions and keywords are used to detect events and topics in a video. The concept of text-based topic detection and tracking for news videos, in which the news clusters are generated based on lexical similarity of news texts was introduced in [14]. Semantics extracted from news clusters for video story boundary detection and search were utilized in [15]. The importance of event text for video search was demonstrated in [24].

The approaches discussed so far belong to a broad category that we define as content based video search, which either uses the low-level visual content or high-level semantic content (or both). While the process of extraction of visual features is usually automatic and domain independent, extracting the semantic content is more complex, because it requires domain knowledge or user interaction or both. The high-level content based approach does not have the limitation of semantic gap. It is based mainly on the attribute information like text annotations and closed captions, which are associated to video manually by human. The process of manual annotation is not only time consuming but also subjective. Moreover, multiple semantic meanings such as metaphorical, hidden or suppressed meanings can be associated with the same video content which makes the process of content description even more complex. For example, a HLF like fire in a video sequence could have different semantic meanings like explosion, forest fire, etc. To overcome the limitations of both the low-level and high-level content based approaches, hybrid video search approaches have been proposed which provide an automatic mapping from low-level features to high-level concepts [25].

CHAPTER III

CONTEXT BASED VIDEO SEARCH

Context based video search approach uses contextual cues to improve search precision. This approach differs from content based approach as it uses story-level contextual cues, instead of (or supplementing) the shot-level visual or semantic content, for video search. The contextual cues intuitively broaden query coverage and facilitate multimodal search.

Commercial video search engines like Google Video and YouTube use text annotations and captions for video search. Repositories of large number of videos are searched using the keywords extracted from captions and text annotations. However, neither the keywords are effectively linked to the video content, nor are they sufficient for to make an effective video search engine. Moreover, the process of building such repositories is user dependant, and relies on the textual content provided by the user while uploading the video. The search process in these search engines is offline as it is dependant on user generated repositories. In the case of news videos, these search engines result in a very disappointing performance as the news videos retrieved are usually old and are presented in an unorganized manner. Moreover, for videos which are uploaded with non-English captions and annotations, the search performance is poor as it relies on translations of the non-English content which often changes the context. All these limitations make these search engines unsuitable for real time and automated video repository generation and search.

A recent product developed by EveryZing (www.everyzing.com) appears to allow the ability to extract and index the full text from any video file, using speech recognition to find spoken words inside videos. Google appears to be experimenting indexing based on audio (www.labs.google.com/gaudi). However, relying completely on speech recognition text may lead to incorrect results, as this technology is still not perfect. An important aspect of any search engine is the ranking of the search results. Commercial search engines like Google use a page ranking scheme that measures the citation importance of a page. Pages with higher ranks are the ones to which a larger number of other pages link with [31]. An outcome of this page ranking scheme is that for a particular query, the same search results are produced, irrespective of the users context of search. However, the context underlying the search for each user may be entirely different. For example a graduate student working on a thesis related to video processing is more likely to search for research papers and articles related to this topic. A query term like video by such a user is more likely to be related to video processing material. On the other hand, the same query for a user who watches a lot of music videos is more likely to be related to music videos rather than research papers on video processing. The context of the user query in the above two cases is very different. The users context of search is also based on the users geographical location. For example, a query term like fire from a user in California is more likely to be related to forest fires in California rather than volcanic fires in Japan. The current search engines provide search results which are same for all the users. Moreover, the search results are fairly static and a query may return the same result for a number of days, until the crawlers update the indexes and the pages are re-ranked. The users objectives for video search on the other hand are dynamic. A system which does a context sensitive ranking of search results can provide much more meaningful results for a user.

We propose a novel context based video clustering and search approach which attempts to make the generation of automated real time video repositories efficient, and also tries to make the process of video browsing and search more meaningful. The

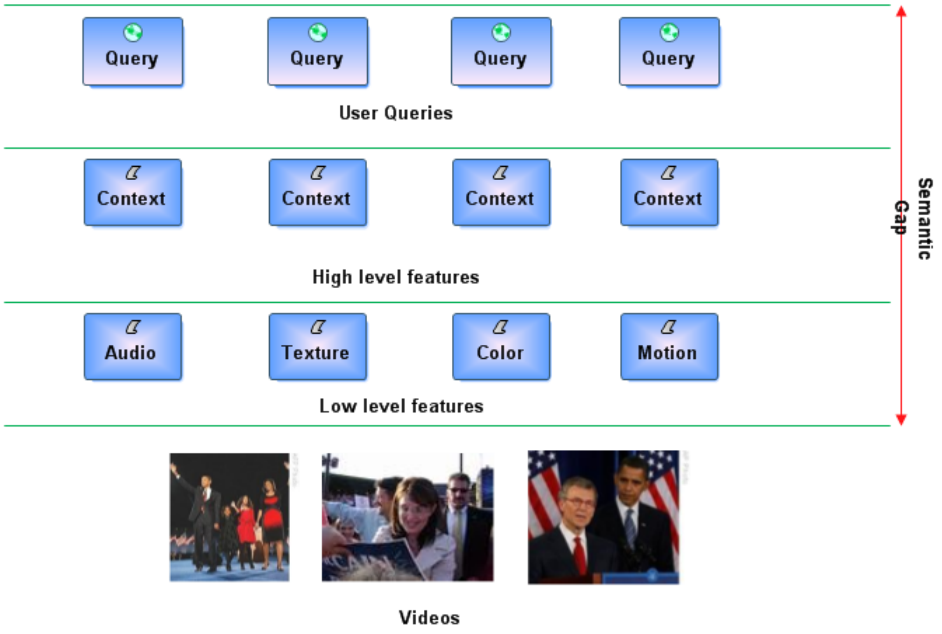


Figure 1: Semantic gap between the low level features and user queries is bridged by the video context.

system is very effective particularly for news video search. Our system is different from the existing context based video search systems for the following reasons:

(1) The news context is derived from a cluster of similar text articles and is used to crawl and cluster videos from different video broadcast sources.

(2) A dynamic mapping from the generated context to the video content is done using the automatic speech recognition (ASR) text, text annotations and closed captions.

(3) A context-sensitive ranking scheme called VideoRank, is used to assign ranks to videos with respect to different contexts.

(4) A query expansion technique is used to enhance the search precision as the users objective of the search may not be clear with the short and imprecise query terms provided.

The above differences are significant in terms of generating good results for video search as the video clustering, ranking and search processes are guided by comprehensive contexts generated from cluster of similar text articles.

There are at least five factors that affect the quality of our search, particularly for news videos:

(1) The context of a news item is dynamic in nature and needs to be updated regularly.

(2) A context derived from multiple news sources is more meaningful, than the one from just a single news source

(3) Clustering of news videos from multiple sources can be made more meaningful based on a context derived from multiple textual news articles

(4) Searching news videos clustered automatically in this manner makes the search process more accurate

(5) Similar news topics and events tend to yield similar videos and thus the clustering of videos can be guided by a comprehensive context generated from several news sources.

CHAPTER IV

FRAMEWORK

Our approach top video search differs from commercial web search engines such as Google, Bing, etc, in that these search engines, crawl and index textual content from HTML pages, PDF documents, etc. Our approach also differs from commercial video search engines such Google Video, YouTube, etc, in that these video search engines rely on the video metadata, which may not provide sufficient information for an effective video search. These video search engines, use keywords from the video descriptions or user tags for search. Moreover, these search engines are not effective for dynamic video content such as news videos. Our approach on the other hand is based on real-time video clustering and search, which is very effective for news videos. The system is completely automated and provides the capability to search, crawl, archive, index, and browse news videos. As a proof of concept, we have created a system called Georgia Tech in the News, for news videos related to Georgia Tech. The framework of the system is described as follows:

4.1 News Clustering

Our system crawls various news sources and internet news clustering services (e.g., Google News or Samachar.com) and extracts all the news items, along with the links to the news articles. A local cache of the news text and links extracted from different news sources is made. The clustered news items are then classified into different context classes (e.g., political, business, weather, sports, entertainment, technology, etc). The idea is that the online news sources broadcasting videos may not have enough text based information to derive a complete context of the news story. Similarly, the online text based news services may not be directly linked to the news video broadcasting websites.

Our system tries to bridge the gap between these two categories of news services. The ASR-text obtained from the video is not always accurate. McCarley and Franz [13] showed that incorrectly recognized speech can often change the context of the news. Therefore, approaches which rely only on extracting the video context from the ASR-text are not accurate. Figure 2 shows the steps involved in news clustering using Google News.

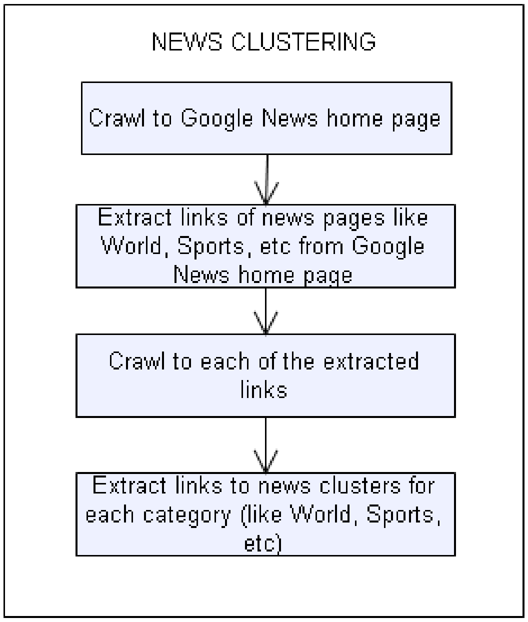

NEWS CLUSTERING
Crawl to Google News home page
Extract links of news pages like World, Sports, etc from Google News home page
Crawl to each of the extracted links
Extract links to news clusters for each category (like World, Sports, etc)

Figure 2: News clustering flow chart

4.2 Context Generation

Our system crawls to various news sources, whose links were extracted from internet news clustering services like Google News and extracts the news text from the web pages. For every news item, the news text from several sources is fetched and cached. The summarization module analyzes the news text from different news sources and creates news summaries. The idea here is to use the news summaries to generate a comprehensive context, which will guide the video clustering and search process.

Thus the news sources that are reporting at the time of video broadcast are used to generate contexts and videos relevant to those contexts are then clustered. The context is dynamic in nature and as newer news items are clustered, the context is updated.

A news context can be divided into two categories, (1) News topic, (2) News event. For example, Presidential elections in US may qualify as a news topic, which gives a broad categorization of news. On the other hand, news like Obama elected as US President, is a news event. The essential difference between these two categories is the lifetime. While a news topic may have a lifetime as long as a year, a news event on the other hand may have a lifetime of only a day. On a higher level, each news context can be classified into a context class. For example, the above news context, along with news topic and event qualify for the political news context class. Such a hierarchical context classification scheme makes the search process more precise. This hierarchical classification scheme is shown in Figure 3. At the highest level are the context classes like political, business, weather, sports, etc, represented as the parent nodes. The next level has news topic contexts, and the third level has the news event contexts. Each news event context has a number of child nodes which are basically the news items clustered in Step II. Every node at each of these four levels is represented by a set of keywords. While the highest level may be represented by a few hundred keywords like rain, temperature, storm, etc., for the weather context class, the lowest level node may have only have a few keywords which are more specific, e.g,. Florida, storm, etc. Thus in this hierarchical structure, a parent node has all the characteristics of a child node.

To create summaries, the system first extracts all the keywords from the news text extracted from different sources. Then a sentence ranking algorithm is used to assign ranks to sentences in the text of different news articles. The following criteria are used for sentence ranking:

1) Location of the sentence in the news article: Generally, sentences which appear in the beginning of the news article contain important content as compared to the sentences toward the end of the article.
2) Number of title keywords in the sentence: This is obtained by the number of matching keywords between the title and sentence.
3) Ratio of number of words in a sentence to the number of keywords: A high ratio generally means that the sentence is important.
4) Sentence length: Shorter sentences having small number of keywords are generally less important.
5) News source ranking: A sentence which comes from a news source with higher rank is given more importance than sentences from lower ranked news sources. The ranking of news sources is done based on criteria like the number of hits, popularity of the news source and the geographical proximity of the news source to the place of origin of a news story.

Two important measures are used for evaluation of the generated summaries. First measure is the compression (C) ratio which is the ratio of the number of words in the summary to the number of words in the original article.

$$C = \frac{\text{Number of words in summary (S)}}{\text{Number of words in original article (O)}}$$

Second measure is the information retention ratio (IR) which is the ratio of the amount of information in the summary to the amount of information in the original article. The number of keywords in the summary and the original article are representative of the amount of information, thus

$$C = \frac{\text{Number of keywords in summary (Ks)}}{\text{Number of keywords in original article (Ko)}}$$

After ranking the sentences, the system finds the union and intersection sets of

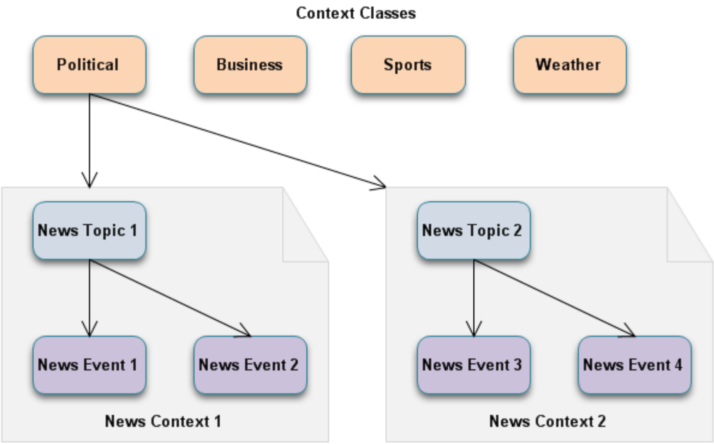


Figure 3: Hierarchical context classification scheme

sentences in news articles from different sources. To generate a short summary, the intersection set is used. The intersection set is formed such that among two similar sentences from different news sources, the one which provides more information (in terms of keywords), and has a higher rank is chosen.

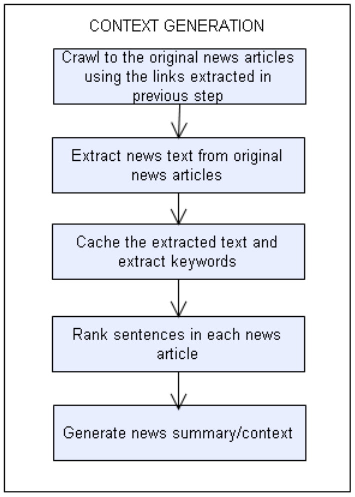

CONTEXT GENERATION
Crawl to the original news articles using the links extracted in previous step
Extract news text from original news articles
Cache the extracted text and extract keywords
Rank sentences in each news article
Generate news summary/context

Figure 4: Context generation flow chart

4.3 Video Crawling

Our system crawls different news sources which broadcast news videos and extracts videos. Using some initial seed links, the video crawlers, crawl web pages containing videos. The crawling process is guided by the kind of videos the system indents to

Table 1: Example of seed URLs for video crawler

Seed URL
http://search.espn.go.com/georgia-tech/video/
http://www.foxnews.com/search-results/search?q=Georgia+Tech&content=Video
http://abcnews.go.com/search?searchtext=georgia%20tech

index. For example, for the Georgia Tech in the News system, the video crawlers crawl web pages containing videos related to Georgia Tech. Example of seed URLs for crawling videos related to Georgia Tech is given in Table-1. These seed URLs are crawled periodically to and the links to the web pages containing videos related to Georgia Tech are extracted. Another alternative way of crawling videos is to send queries to the commercial video search engines such as Google Video, YouTube, etc and use the search results as the seed links. However, this approach is not suitable for dynamic content such as news videos as these commercial video search engines take long time to crawl and index new videos. The video crawling process involves extracting the links of web pages in which the videos are embedded. The next step is to extract the videos from these web pages.

4.4 Video Extraction

Once the video crawlers, crawl and extract the links of web pages containing videos, the video extraction engine, extracts and downloads the videos to a local disk. Most video broadcasting websites use provide videos through streaming. The videos are embedded in web pages using different video players which use different formats. The commonly used formats are Flash and MP4. The links extracted by video crawlers are not the direct links to the videos, since the videos are embedded in the web pages. Therefore to get the direct link to the videos, our system uses streamsiff, which is a utility that sniffs network traffic for stream URLs. This utility detects

RTSP, and other video streaming protocols, and performs a back-trace on HTTP traffic to detect the video URLs and Flash video files. The direct links to the video files obtained from stream sniffing are then used to extract and download the videos. For this, a command line base downloader utility is used, which downloads the videos to a local disk. Since the videos files are large in size, typically few 100 Megabytes on average, downloading videos sequentially takes a significant amount of time. To minimize the download time the system opens parallel downloads streams.

4.5 Video Metadata Extraction

The websites publishing videos provide some metadata related to videos. The amount of metadata varies from one site to other as no particular standard is followed. For example, YouTube provides meta information such as the video title, description, user tags, author, submission time, video length, recording date, user comments, etc. Such meta information extracted from the web pages containing videos is valuable for video search.

4.6 Video Transcription

The video transcription engine uses speech recognition technology to transcribe the videos. The videos extracted by the video extraction engine are in different formats such as FLV, MP4, etc. For transcription, the audio channel of the video is extracted and then speech recognition tools are used to convert the audio to text. To extract the audio channel from the video, we use an open source library called ffmpeg, which has implementation supporting most codec. Using ffmpeg, the audio files in WAV format are obtained from the videos. These WAV audio files are then used as input to standard speech recognition tools for transcription. We use a speech recognition tool from Dragon for this purpose.

4.7 Video Clustering

The video clustering process is guided by the news contexts generated in the previous step. The video clustering is done in two steps. In the first step, the textual content surrounding these videos, including the captions and annotations is mapped to the contexts generated in the previous step. Based on this mapping the system classifies the extracted news videos into the news contexts. In the second step automatic speech recognition (ASR) is performed on the clustered videos and a more precise mapping is done to the news contexts. The two step approach makes the system efficient for real time video clustering, as time consuming process of speech recognition is not involved in the initial clustering of the videos. The system does not rely completely on the ASR-text, as it is not always accurate. However, the ASR-text is useful to extract some keywords from the video speech, which are used for a more precise mapping to the news contexts. This precise mapping is used to perform a re-clustering of videos to different contexts, in case there is an error in the initial clustering which relies only on the text annotations and captions. Here an MPEG-7 based framework is used for video description. MPEG-7 descriptors like AudioVisualSegment, MediaTime, MediaUri, MediaLocator, TextAnnotation, KeywordAnnotation, etc are used to capture the metadata in an XML based format.

For classifying the videos into different news contexts a probabilistic support vector machine (PSVM) with pairwise coupling (PWC) is used. Let V be a set of n videos $V = \{v_1, v_2, v_3, ..., v_n\}$ and C be a set of k contexts $C = \{c_1, c_2, c_3, ..., c_k\}$. A probabilistic model for classification of videos into contexts will select a context $c_i \in C$ for video $v \in V$ with probability $p(c_i|v)$. The conditional probability $p(c_i|v)$ is estimated using probabilistic support vector machine (PSVM) with pairwise coupling (PWC). The classification procedure is described below.

Suppose that we are given l training vectors (videos) $x_i(1 < i < l)$, where x_i is a feature vector in n dimensional feature space and y_i is the class label of x_i such that

$$y_i = \begin{cases} 1 & \text{if } x_i \text{ in class } 1 \\ -1 & \text{if } x_i \text{ in class } 2 \end{cases}$$

A standard SVM finds a hyperplane $w^T x + b = 0$ which correctly, separates the training vectors and has a maximum margin which is the distance between two, hyperplanes $w^T x + b = 1$ and -1. The optimal hyperplane with maximum margin can be obtained by solving the following quadratic programming problem,

$$\min_{w,b} \frac{1}{2}||w||^2 + C\sum_{i=1}^{l} \xi_i \text{ , subject to } y_i(w.x_i + b) > 1 - \xi_i, \xi_i > 0, (1 < i < l)$$

where C is the constant and ξ_i is a slack variable for the non-separable case.

The optimal hyperplane is given as,

$$f(x) = sign\left(C\sum_{i=1}^{l} \alpha_i y_i K(x_i, x) + b\right)$$

where α_i is the Lagrange multiple, and $K(x_i, x)$ is a kernel function. The SVM calculates similarity between two arguments x_i and x. A standard SVM is a two-class classifier where the outcome y is 1 or 1. The classifier predicts class 1 if $w^T x + b > 0$ and class 2 otherwise. An extension to SVM called the probabilistic SVM can produce a posteriori class probabilities $P(class|input)$. A sigmoid model maps the binary SVM scores to posterior probabilities, where the probability of membership in class y, $y \in \{+1, 1\}$ is given by

$$p(y|x) = \frac{1}{1 + exp(Af(x) + B)}$$

where $f(x)$ is the output of the SVM decision function and A and B are the parameters of the sigmoid function. A and B are found by minimizing the class entropy of the training data. First the SVM is trained and then the parameters of the sigmoid function are trained to map the SVM outputs into probabilities [32].

To classify the video in the set $V = \{v_1, v_2, v_3, ..., v_n\}$ into contexts in the set $C = \{c_1, c_2, c_3, ..., c_k\}$ a pairwise coupling procedure is used [33]. Using the probabilistic SVM, we write the posterior probability of video v belonging to context c_i, given that v belongs to either c_i or c_j as the pairwise probability $r_{ij} = p(c_i|r \in rc_i \cup c_j)$. For classification into k contexts the pairwise coupling method trains $k(k-1)/2$ SVM classifiers. Going through $k(k-1)/2$ SVM classifiers a pairwise probabilities matrix(PPM) is obtained. To couple the PPM into a common set of posterior probabilities $p(c_i|v)$, [33] used the auxiliary variables

$$u_{ij} = \frac{p_i}{p_i + p_j}$$

and found p_i's such that u_{ij}'s are close to the r_{ij}'s. Hastie and Tibshirani [33] proposed to minimize the KullbackLeibler distance between the u_{ij} and r_{ij} as the closeness criterion.

$$l(p) = \sum_{i<j} n_{ij} \left[log\frac{r_{ij}}{u_{ij}} + (1 - r_{ij}) log\frac{(1 - r_{ij})}{(1 + u_{ij})} \right]$$

where n_{ij} are the number of observations in the training set and p_i's are found to minimize the function $l(p)$.

The probability $p(c_i|v)$ can be used to score context c among possible contexts for video v. The videos are also classified into context classes, like political, business, weather, sports, entertainment, technology, etc. This classification process is guided by context class lexicons. For each context class there is a separate lexicon which contains the frequently used keywords. For example, in a weather news, keywords like rain, storm, temperature are common. Other information like the video broadcast time, source and news event date is also tagged with the video, to help in the search process. The system makes a news video repository, which allows users to search news videos. The popularity of video formats like FLV and MP4 have made embedding and extraction of videos from the original sources easier. Our system differs from video sharing websites like YouTube, in the sense that such websites rely on the users to upload videos manually and attach captions and annotation for efficient search. Our system is completely automated and the videos are clustered and tagged without any user intervention.

4.8 Video Indexing and Ranking

Video indexing and ranking is an important step for efficient browsing and search of videos. After clustering the videos our system assigns relevance weights to the videos with respect to a context and calculates the VideoRank. VideoRank indicates the likelihood of the presence of a context in a video. This ranking scheme is different from the page ranking schemes used by search engines like Google, where pages with more citations are ranked higher. Such a ranking scheme does not work well for video search in a particular context as a page may have a number of embedded videos, which need to be ranked individually according to their relevance to a context. Our approach to video ranking is based on the context. A number of criteria are used for calculating the VideoRank, like the number of matching keywords between the news context and the video metadata information. The news source ranking is also taken into account for the process of video ranking. The news source ranking is based on criteria like the number of hits, popularity of the news source and the geographical proximity of the news source to the place of origin of a news story. As in the case of context generation, the video ranking process is also dynamic in nature. As the system constantly clusters videos, the rankings also keep changing. Other criteria like

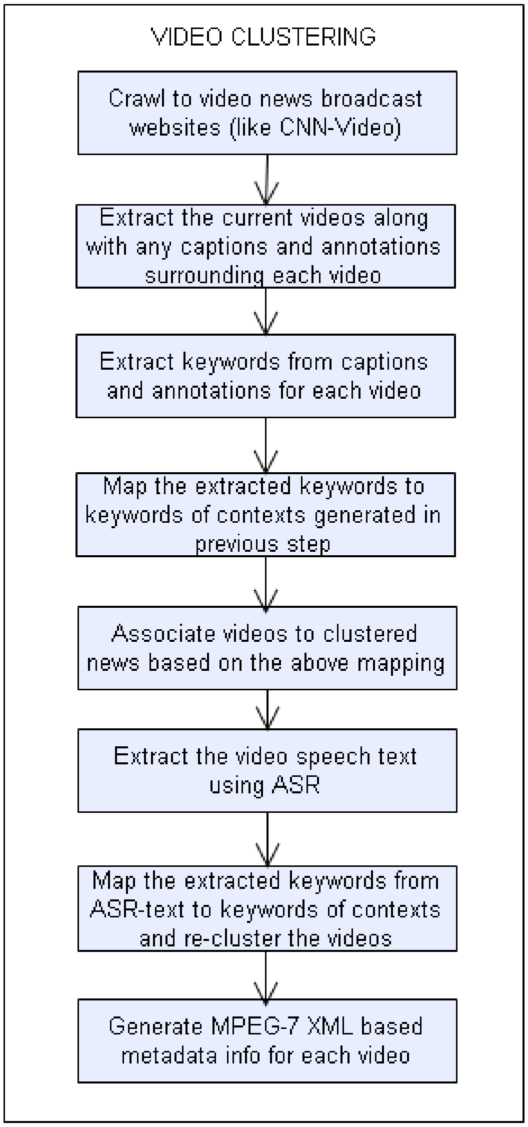

VIDEO CLUSTERING
Crawl to video news broadcast websites (like CNN-Video)
Extract the current videos along with any captions and annotations surrounding each video
Extract keywords from captions and annotations for each video
Map the extracted keywords to keywords of contexts generated in previous step
Associate videos to clustered news based on the above mapping
Extract the video speech text using ASR
Map the extracted keywords from ASR-text to keywords of contexts and re-cluster the videos
Generate MPEG-7 XML based metadata info for each video

Figure 5: Video clustering flow chart

the time of broadcast of video and news event date can also be taken into account. Thus newer videos matching a particular context may be ranked higher than the older videos.

The VideoRank of a video v is defined as,

$$VR(v) = R(c) + R(n)$$

where $R(c)$ is the relevance rank of the video with respect to a context c and $R(n)$

is the rank of news source n. Let C be a set of N contexts $C = \{c_1, c_2, c_3, ..., c_N\}$, where context ci is charactered by a set of M keywords $K = \{k_1, k_2, k_3, ..., k_M\}$. Every context may have a number of keywords which may be common with other contexts. A keyword which occurs in many contexts is not a good discriminator, and should be given less weight than one which occurs in few contexts. To find the discriminating power of a keyword for a context, we calculate the inverse document frequency,

$$IDF(k_i) = log(N)/n_i$$

where N is the total number of contexts and n_i is the number of contexts in which the keyword k_i occurs. The relevance rank of the video with respect to a context c is found as defined as follows,

$$R(c) = \Sigma IDF(y_i)TF(y_i)$$

where y_i are the matching keywords between the context c and the video metadata information (which includes the ASR text, captions and annotations). $TF(y_i)$ is the term frequency of the keyword y_i i.e., the number of times the keyword occurs in the video metadata information. A user query may have different contexts with respect to different domains. For example, if the user has sports in mind while searching for a term like ”videos”, then he is clearly interested in sports videos and not music videos. Our system not only ranks the results based on the context but also provides a clear separation of different domains of search like sports, politics, weather, etc. A domain classification module classifies different search results into different domains. A domain relevance rank is computed for each search result and based on this rank the search result is classified to a particular domain. There may be a case where the relevance rank for a search result is almost the same for two or more domains. In this case the result is classified to all the domains for which the relevance rank is greater than a threshold. Ranking of the search result is then done separately in each domain as compared to the other results.

4.9 Video Search

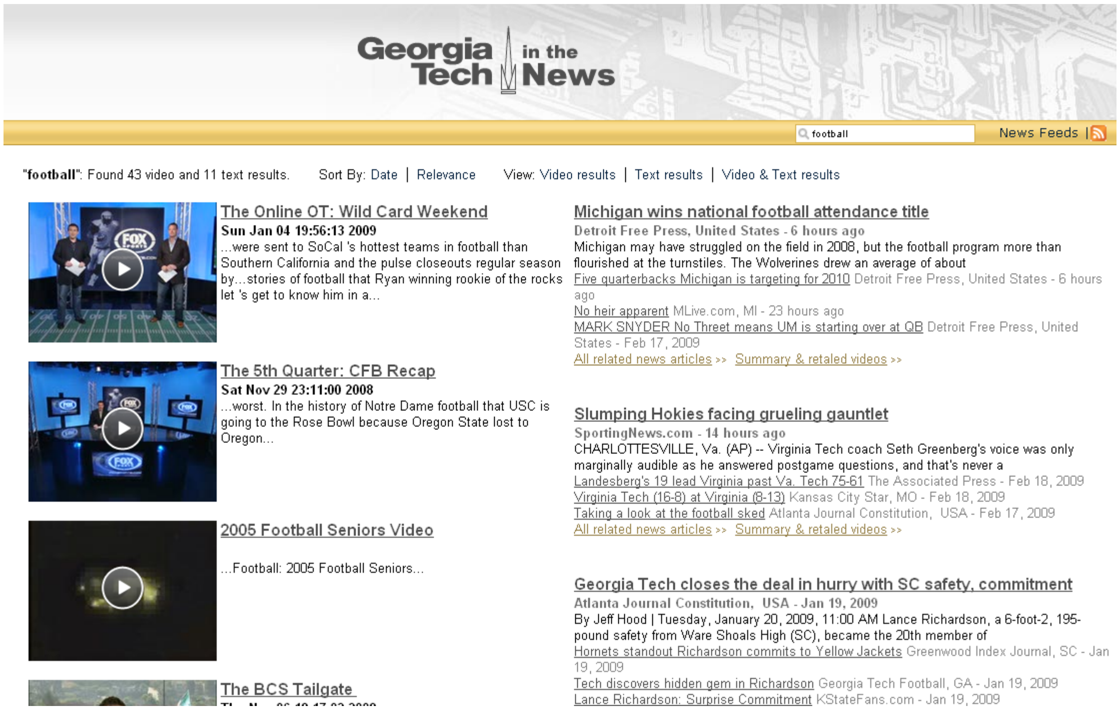
Queries for video search that are short may not contain enough information to map it to one of the contexts generated in Step III. Our system uses a query expansion technique to enhance contextual information of the query which can then be mapped more efficiently to the contexts generated in Step III, thus enhancing the recall and improving the search precision. [11],[12] have shown the usefulness of this query expansion technique.

Given an input query, first the query terms are stemmed using the Porter Stemming Algorithm to reduce the query terms to their base or root form. Then the stop words such as a, an, the, etc from are removed from the query. The normalized query is then expanded using a query expansion algorithm. In query expansion the seed query is reformulated in order to increase the precision of recall. The idea is to analyze the query terms and find other similar terms or keywords which have a high correlation with the query terms. The steps involved in query processing and expansion are as follows:

1) Query terms are reduced to their root form using Porter Stemming Algorithm.
2) Stop words are removed from the query.
3) Query is expanded to include synonyms of the query terms.
4) Current news contexts are searched for matching terms as in the query and keywords from the context are used for further expanding the query.

For example, consider a query which has only one term elections. This single term cannot give any contextual information. However based on the current news clusters and the generated contexts, the query can be expanded such that keywords among all the generated contexts which have a high correlation with the query term are added to the query. So if there is a current news cluster on Presidential elections in US, then the system will attach keywords like President, US, etc to the original query. Our approach to query expansion in novel in the sense that it takes into consideration the contexts generated from the current news clusters. As queries are sensitive to time, the query expansion process is dynamic in nature and depends on the current contexts.

To retrieve the most relevant videos for a given query, the system uses the expanded query and the video contexts. Using these two techniques, the total recall as well as the precision of recall is greatly enhanced. To illustrate this by an example, consider a query such as football to the Georgia Tech in the News system. Although the system crawls and clusters videos related to football daily from many video broadcasting websites, there may be many videos for which the metadata does not have the word football. However, there might be terms such as Georgia Tech Yellow Jackets, associated with some videos, which is the name used by the Georgia Tech football team. Since the video clustering and search process is guided by the context generated from clusters of news articles, as described in section 4.2, the terms such as football will be associated with the videos which do not have that term in the metadata but have other terms which are highly correlated to the search term. This increases the total recall for queries, and the videos that could not be retrieved merely by keyword-based search, can be retrieved using the video contexts. Therefore for each term in the expanded query, the system will return the news clusters and the related videos which have that term as a part of the context. Figure 6 shows the results for the query - football.



"**football**": Found 43 video and 11 text results. Sort By: Date | Relevance View: Video results | Text results | Video & Text results

The Online OT: Wild Card Weekend
Sun Jan 04 19:56:13 2009
...were sent to SoCal 's hottest teams in football than Southern California and the pulse closeouts regular season by...stories of football that Ryan winning rookie of the rocks let 's get to know him in a...

The 5th Quarter: CFB Recap
Sat Nov 29 23:11:00 2008
...worst. In the history of Notre Dame football that USC is going to the Rose Bowl because Oregon State lost to Oregon...

2005 Football Seniors Video

...Football: 2005 Football Seniors...

The BCS Tailgate

Michigan wins national football attendance title
Detroit Free Press, United States - 6 hours ago
Michigan may have struggled on the field in 2008, but the football program more than flourished at the turnstiles. The Wolverines drew an average of about
Five quarterbacks Michigan is targeting for 2010 Detroit Free Press, United States - 6 hours ago
No heir apparent MLive.com, MI - 23 hours ago
MARK SNYDER No Threet means UM is starting over at QB Detroit Free Press, United States - Feb 17, 2009
All related news articles >> Summary & retaled videos >>

Slumping Hokies facing grueling gauntlet
SportingNews.com - 14 hours ago
CHARLOTTESVILLE, Va. (AP) -- Virginia Tech coach Seth Greenberg's voice was only marginally audible as he answered postgame questions, and that's never a
Landesberg's 19 lead Virginia past Va. Tech 75-61 The Associated Press - Feb 18, 2009
Virginia Tech (16-8) at Virginia (8-13) Kansas City Star, MO - Feb 18, 2009
Taking a look at the football sked Atlanta Journal Constitution, USA - Feb 17, 2009
All related news articles >> Summary & retaled videos >>

Georgia Tech closes the deal in hurry with SC safety, commitment
Atlanta Journal Constitution, USA - Jan 19, 2009
By Jeff Hood | Tuesday, January 20, 2009, 11:00 AM Lance Richardson, a 6-foot-2, 195-pound safety from Ware Shoals High (SC), became the 20th member of
Hornets standout Richardson commits to Yellow Jackets Greenwood Index Journal, SC - Jan 19, 2009
Tech discovers hidden gem in Richardson Georgia Tech Football, GA - Jan 19, 2009
Lance Richardson: Surprise Commitment KStateFans.com - Jan 19, 2009

Figure 6: Results of video search for the query - football

4.10 Query Lifecycle

Figure 7 shows the lifecycle of a user query. The user query is first sent to web server which sends it to the query expansion agent. The expanded query is then sent to the index servers. Index servers have indexes of video meta-data including the video ASR text and the video context. The query is then sent to the context servicing and VideoRank agent which ranks the videos based on the context as described in section V. The content delivery servers retrieve the videos and generate the content describing the video search result. The video search results are then returned to the user.

4.11 System Architecture

Figure 8 shows the system architecture. The news crawler crawls various news sources and internet news clustering services and extracts all the news items, along with the links to the news articles. The news text extraction module then crawls to the original

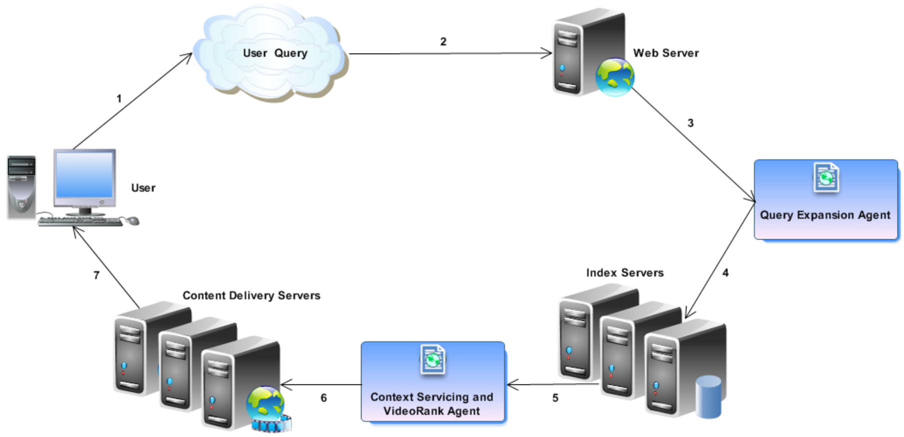


Figure 7: Video query lifecycle

news articles and extracts the news text. The news classification module classifies the news into different context classes like political, business, weather, sports, etc. The news summarization and context generation module analyses the news text from different news sources and generates news summaries and contexts. Guided by the generated contexts the video crawler crawls different news video broadcast websites and extracts videos.

The video clustering module then clusters videos in two steps. In the first step, the textual content surrounding the videos, including the captions and annotations is mapped to the news contexts. Based on this mapping the module classifies the extracted news videos into the news contexts. In the second step automatic speech recognition (ASR) is performed on the clustered videos and a more precise mapping is done to the news contexts. The indexing module indexes video meta-data including the video ASR text and the video context.

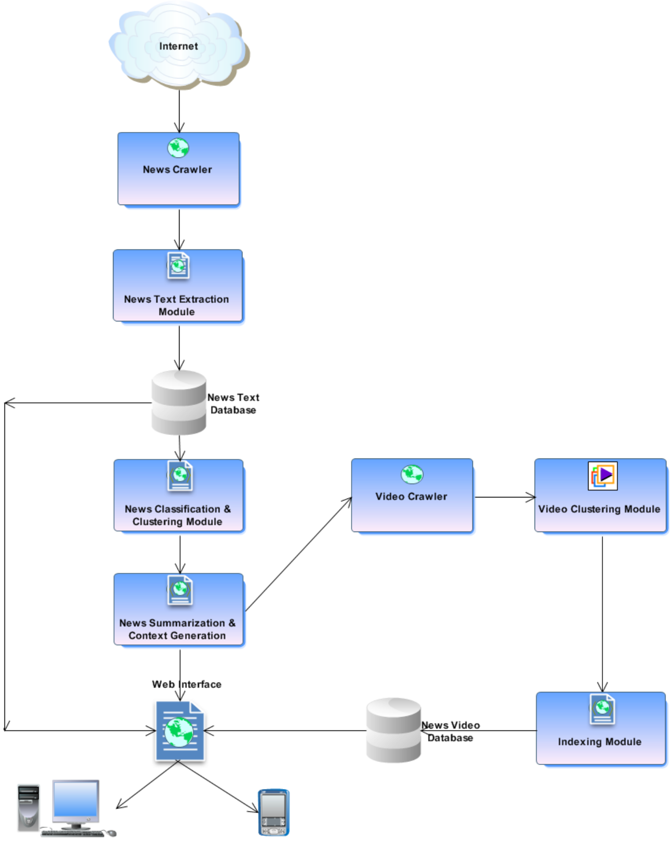
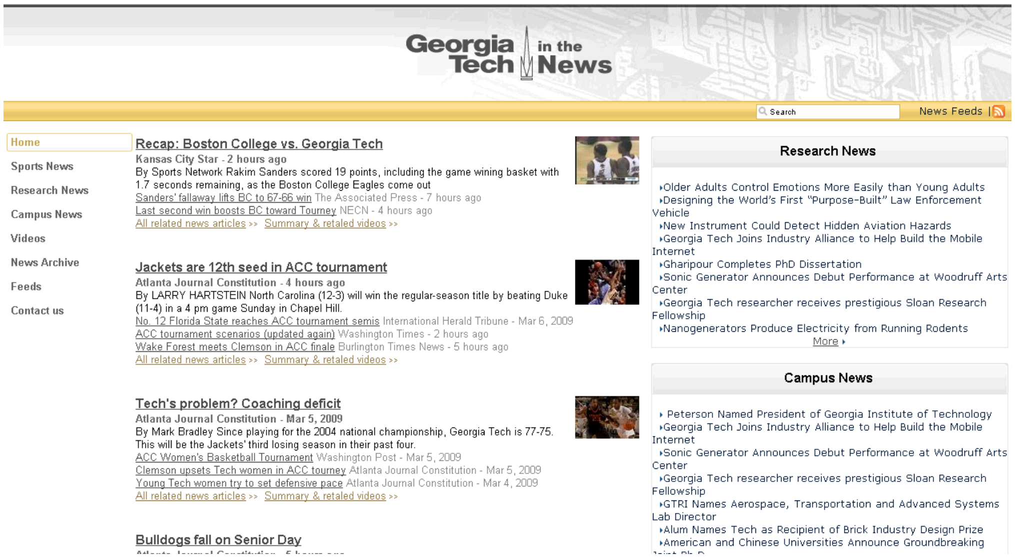


Figure 8: System Architecture

CHAPTER V

USER INTERFACE

Commercial news services like Google News provide links to the news stories clustered from several sources. Links to the original news sources are presented and the user has to visit different news articles to get a comprehensive view of the news story. Such an interface may present links to hundreds of news articles for a particular news story which is generally overwhelming for most users. Video search engines like YouTube, on the other hand provide lists of videos arranged in order of relevance. This interface again overwhelms the users with hundreds of videos, most of which may not be relevant to the user in a particular context. Our system overcomes the limitations of both the commercial news services like Google News and video search engines like YouTube.



Recap: Boston College vs. Georgia Tech
Kansas City Star - 2 hours ago
By Sports Network Rakim Sanders scored 19 points, including the game wining basket with 1.7 seconds remaining, as the Boston College Eagles come out
Sanders' fallaway lifts BC to 67-66 win The Associated Press - 7 hours ago
Last second win boosts BC toward Tourney NECN - 4 hours ago
All related news articles >> Summary & retaled videos >>

Jackets are 12th seed in ACC tournament
Atlanta Journal Constitution - 4 hours ago
By LARRY HARTSTEIN North Carolina (12-3) will win the regular-season title by beating Duke (11-4) in a 4 pm game Sunday in Chapel Hill.
No. 12 Florida State reaches ACC tournament semis International Herald Tribune - Mar 6, 2009
ACC tournament scenarios (updated again) Washington Times - 2 hours ago
Wake Forest meets Clemson in ACC finale Burlington Times News - 5 hours ago
All related news articles >> Summary & retaled videos >>

Tech's problem? Coaching deficit
Atlanta Journal Constitution - Mar 5, 2009
By Mark Bradley Since playing for the 2004 national championship, Georgia Tech is 77-75. This will be the Jackets' third losing season in their past four.
ACC Women's Basketball Tournament Washington Post - Mar 5, 2009
Clemson upsets Tech women in ACC tourney Atlanta Journal Constitution - Mar 5, 2009
Young Tech women try to set defensive pace Atlanta Journal Constitution - Mar 4, 2009
All related news articles >> Summary & retaled videos >>

Bulldogs fall on Senior Day

Research News

- Older Adults Control Emotions More Easily than Young Adults
- Designing the World's First "Purpose-Built" Law Enforcement Vehicle
- New Instrument Could Detect Hidden Aviation Hazards
- Georgia Tech Joins Industry Alliance to Help Build the Mobile Internet
- Gharipour Completes PhD Dissertation
- Sonic Generator Announces Debut Performance at Woodruff Arts Center
- Georgia Tech researcher receives prestigious Sloan Research Fellowship
- Nanogenerators Produce Electricity from Running Rodents

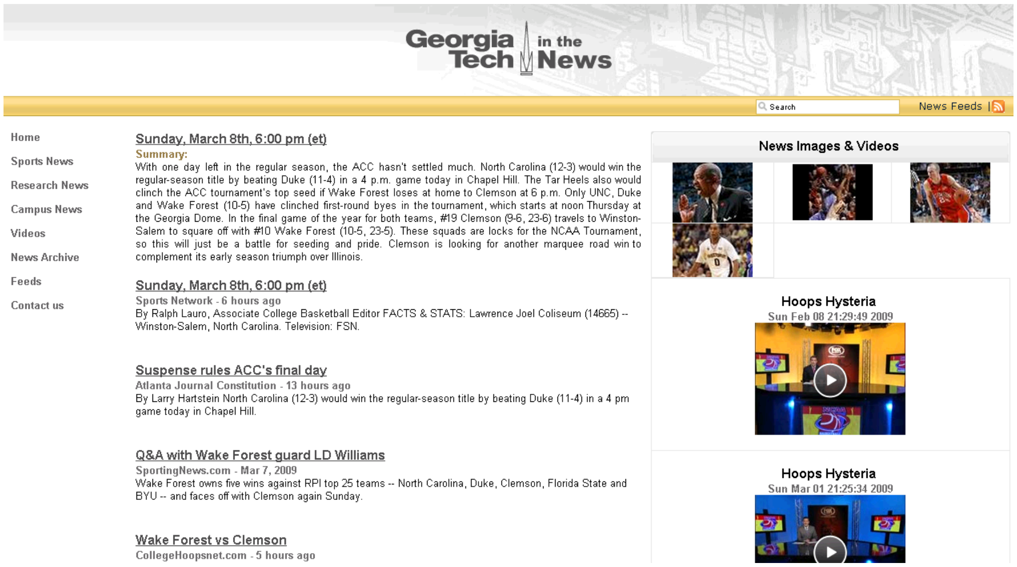
More

Campus News

- Peterson Named President of Georgia Institute of Technology
- Georgia Tech Joins Industry Alliance to Help Build the Mobile Internet
- Sonic Generator Announces Debut Performance at Woodruff Arts Center
- Georgia Tech researcher receives prestigious Sloan Research Fellowship
- GTRI Names Aerospace, Transportation and Advanced Systems Lab Director
- Alum Names Tech as Recipient of Brick Industry Design Prize
- American and Chinese Universities Announce Groundbreaking

Figure 9: Screen shot of home page, showing the news clusters

A novel user interface is provided which not only gives the links to the original news articles but also provides the news summaries, related news videos and images



Sunday, March 8th, 6:00 pm (et)
Summary:
With one day left in the regular season, the ACC hasn't settled much. North Carolina (12-3) would win the regular-season title by beating Duke (11-4) in a 4 p.m. game today in Chapel Hill. The Tar Heels also would clinch the ACC tournament's top seed if Wake Forest loses at home to Clemson at 6 p.m. Only UNC, Duke and Wake Forest (10-5) have clinched first-round byes in the tournament, which starts at noon Thursday at the Georgia Dome. In the final game of the year for both teams, #19 Clemson (9-6, 23-6) travels to Winston-Salem to square off with #10 Wake Forest (10-5, 23-5). These squads are locks for the NCAA Tournament, so this will just be a battle for seeding and pride. Clemson is looking for another marquee road win to complement its early season triumph over Illinois.

Sunday, March 8th, 6:00 pm (et)
Sports Network - 6 hours ago
By Ralph Lauro, Associate College Basketball Editor FACTS & STATS: Lawrence Joel Coliseum (14665) -- Winston-Salem, North Carolina. Television: FSN.

Suspense rules ACC's final day
Atlanta Journal Constitution - 13 hours ago
By Larry Hartstein North Carolina (12-3) would win the regular-season title by beating Duke (11-4) in a 4 pm game today in Chapel Hill.

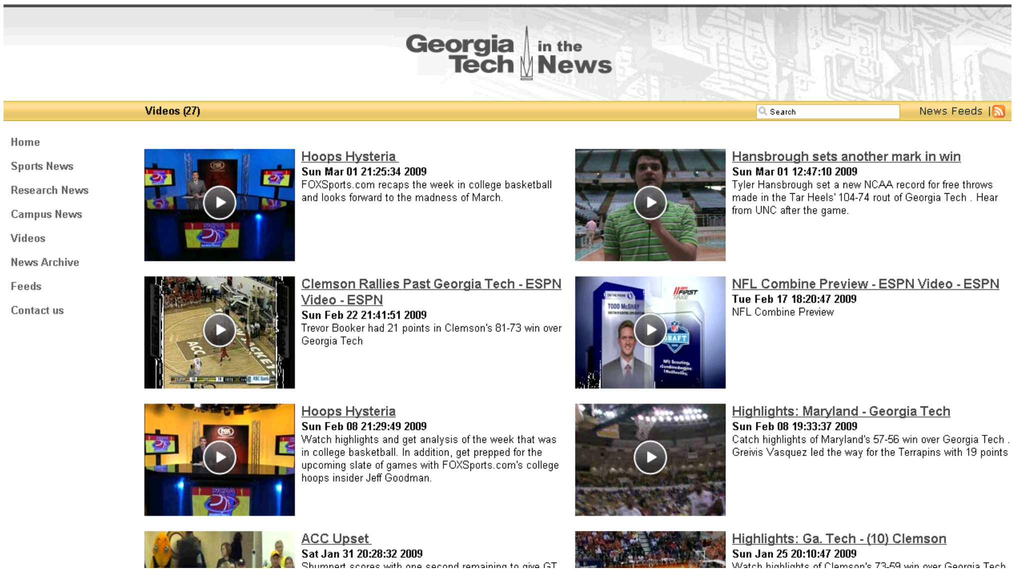
Q&A with Wake Forest guard LD Williams
SportingNews.com - Mar 7, 2009
Wake Forest owns five wins against RPI top 25 teams -- North Carolina, Duke, Clemson, Florida State and BYU -- and faces off with Clemson again Sunday.

Wake Forest vs Clemson
CollegeHoopsnet.com - 5 hours ago

Hoops Hysteria
Sun Feb 08 21:29:49 2009

Hoops Hysteria
Sun Mar 01 21:25:34 2009

Figure 10: News summary page with related videos and images



Hoops Hysteria
Sun Mar 01 21:25:34 2009
FOXSports.com recaps the week in college basketball and looks forward to the madness of March.

Hansbrough sets another mark in win
Sun Mar 01 12:47:10 2009
Tyler Hansbrough set a new NCAA record for free throws made in the Tar Heels' 104-74 rout of Georgia Tech . Hear from UNC after the game.

Clemson Rallies Past Georgia Tech - ESPN Video - ESPN
Sun Feb 22 21:41:51 2009
Trevor Booker had 21 points in Clemson's 81-73 win over Georgia Tech

NFL Combine Preview - ESPN Video - ESPN
Tue Feb 17 18:20:47 2009
NFL Combine Preview

Hoops Hysteria
Sun Feb 08 21:29:49 2009
Watch highlights and get analysis of the week that was in college basketball. In addition, get prepped for the upcoming slate of games with FOXSports.com's college hoops insider Jeff Goodman.

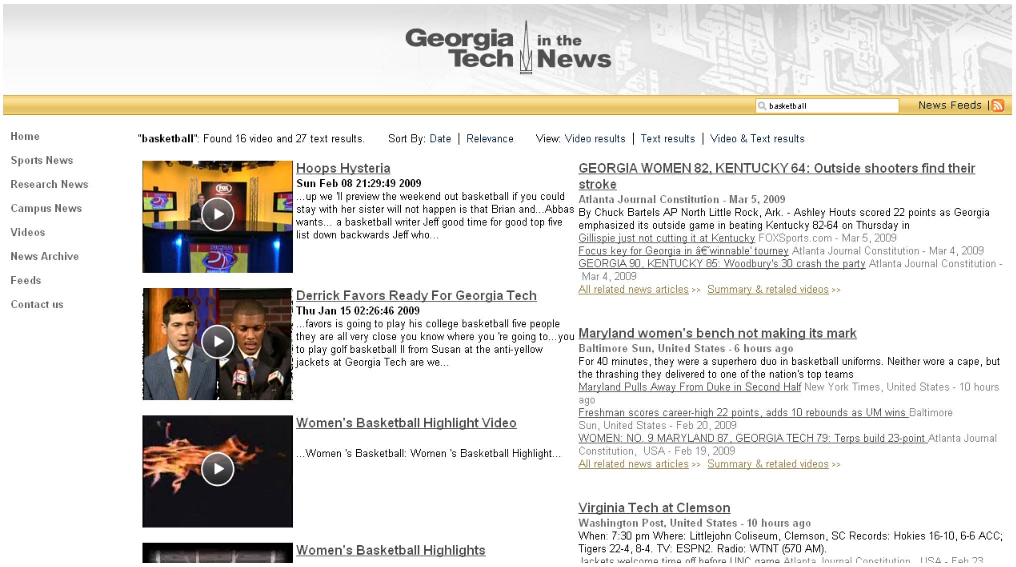
Highlights: Maryland - Georgia Tech
Sun Feb 08 19:33:37 2009
Catch highlights of Maryland's 57-56 win over Georgia Tech . Greivis Vasquez led the way for the Terrapins with 19 points

ACC Upset
Sat Jan 31 20:28:32 2009
Shumpert scores with one second remaining to give GT

Highlights: Ga. Tech - (10) Clemson
Sun Jan 25 20:10:47 2009
Watch highlights of Clemson's 73-59 win over Georgia Tech

Figure 11: Screenshot of news videos page showing the videos along with the related metadata

all at one place. This makes the process of news search more interesting as a user can read a brief summary of the news and watch related news videos at the same time. Due to the dynamic nature of the content on the web, a user may be interested in getting automated updates for a query. For example, a user who is interested in videos of a sports tournament or weather related videos may be interested in getting



"**basketball**": Found 16 video and 27 text results. Sort By: Date | Relevance View: Video results | Text results | Video & Text results

Hoops Hysteria
Sun Feb 08 21:29:49 2009
...up we 'll preview the weekend out basketball if you could stay with her sister will not happen is that Brian and...Abbas wants... a basketball writer Jeff good time for good top five list down backwards Jeff who...

Derrick Favors Ready For Georgia Tech
Thu Jan 15 02:26:46 2009
...favors is going to play his college basketball five people they are all very close you know where you 're going to...you to play golf basketball ll from Susan at the anti-yellow jackets at Georgia Tech are we...

Women's Basketball Highlight Video
...Women 's Basketball: Women 's Basketball Highlight...

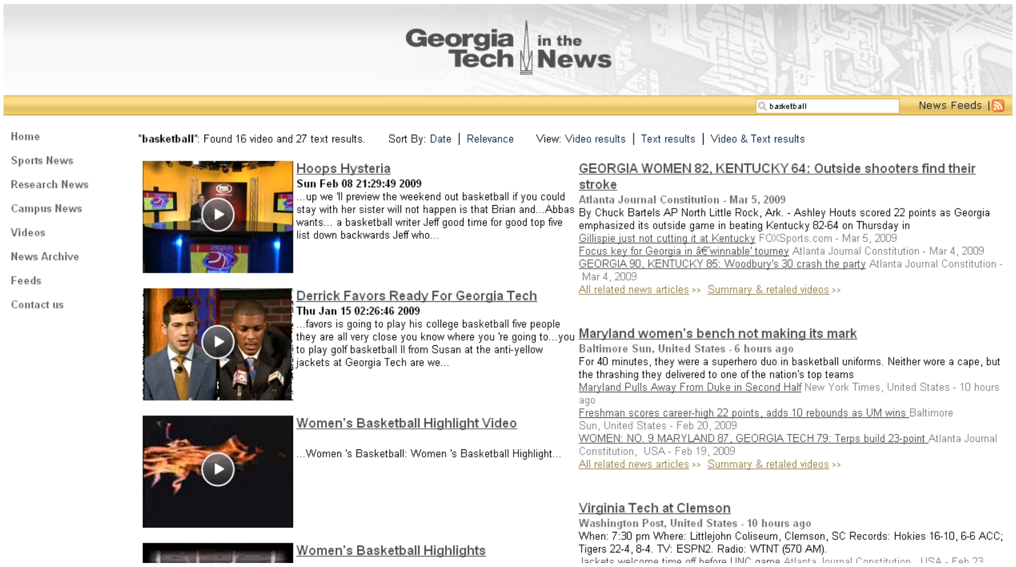
Women's Basketball Highlights

GEORGIA WOMEN 82, KENTUCKY 64: Outside shooters find their stroke
Atlanta Journal Constitution - Mar 5, 2009
By Chuck Bartels AP North Little Rock, Ark. - Ashley Houts scored 22 points as Georgia emphasized its outside game in beating Kentucky 82-64 on Thursday in
Gillispie just not cutting it at Kentucky FOXSports.com - Mar 5, 2009
Focus key for Georgia in â€˜winnable' tourney Atlanta Journal Constitution - Mar 4, 2009
GEORGIA 90, KENTUCKY 85: Woodbury's 30 crash the party Atlanta Journal Constitution - Mar 4, 2009
All related news articles >> Summary & retaled videos >>

Maryland women's bench not making its mark
Baltimore Sun, United States - 6 hours ago
For 40 minutes, they were a superhero duo in basketball uniforms. Neither wore a cape, but the thrashing they delivered to one of the nation's top teams
Maryland Pulls Away From Duke in Second Half New York Times, United States - 10 hours ago
Freshman scores career-high 22 points, adds 10 rebounds as UM wins Baltimore Sun, United States - Feb 20, 2009
WOMEN: NO. 9 MARYLAND 87, GEORGIA TECH 79: Terps build 23-point Atlanta Journal Constitution, USA - Feb 19, 2009
All related news articles >> Summary & retaled videos >>

Virginia Tech at Clemson
Washington Post, United States - 10 hours ago
When: 7:30 pm Where: Littlejohn Coliseum, Clemson, SC Records: Hokies 16-10, 6-6 ACC; Tigers 22-4, 8-4. TV: ESPN2. Radio: WTNT (570 AM).
Jackets welcome time off before UNC game Atlanta Journal Constitution, USA - Feb 23

Figure 12: Video search results for a query basketball



"**basketball**": Found 16 video and 27 text results. Sort By: Date | Relevance View: Video results | Text results | Video & Text results

Hoops Hysteria
Sun Feb 08 21:29:49 2009
...up we 'll preview the weekend out basketball if you could stay with her sister will not happen is that Brian and...Abbas wants... a basketball writer Jeff good time for good top five list down backwards Jeff who...

Derrick Favors Ready For Georgia Tech
Thu Jan 15 02:26:46 2009
...favors is going to play his college basketball five people they are all very close you know where you 're going to...you to play golf basketball ll from Susan at the anti-yellow jackets at Georgia Tech are we...

Women's Basketball Highlight Video
...Women 's Basketball: Women 's Basketball Highlight...

Women's Basketball Highlights

GEORGIA WOMEN 82, KENTUCKY 64: Outside shooters find their stroke
Atlanta Journal Constitution - Mar 5, 2009
By Chuck Bartels AP North Little Rock, Ark. - Ashley Houts scored 22 points as Georgia emphasized its outside game in beating Kentucky 82-64 on Thursday in
Gillispie just not cutting it at Kentucky FOXSports.com - Mar 5, 2009
Focus key for Georgia in â€˜winnable' tourney Atlanta Journal Constitution - Mar 4, 2009
GEORGIA 90, KENTUCKY 85: Woodbury's 30 crash the party Atlanta Journal Constitution - Mar 4, 2009
All related news articles >> Summary & retaled videos >>

Maryland women's bench not making its mark
Baltimore Sun, United States - 6 hours ago
For 40 minutes, they were a superhero duo in basketball uniforms. Neither wore a cape, but the thrashing they delivered to one of the nation's top teams
Maryland Pulls Away From Duke in Second Half New York Times, United States - 10 hours ago
Freshman scores career-high 22 points, adds 10 rebounds as UM wins Baltimore Sun, United States - Feb 20, 2009
WOMEN: NO. 9 MARYLAND 87, GEORGIA TECH 79: Terps build 23-point Atlanta Journal Constitution, USA - Feb 19, 2009
All related news articles >> Summary & retaled videos >>

Virginia Tech at Clemson
Washington Post, United States - 10 hours ago
When: 7:30 pm Where: Littlejohn Coliseum, Clemson, SC Records: Hokies 16-10, 6-6 ACC; Tigers 22-4, 8-4. TV: ESPN2. Radio: WTNT (570 AM).
Jackets welcome time off before UNC game Atlanta Journal Constitution, USA - Feb 23

Figure 13: Video summaries using keyframes

automated updates whenever new videos are available. Our system provides this feature by the creation of custom feeds for a particular query. As new videos are crawled, the system can send updates to an interested user.

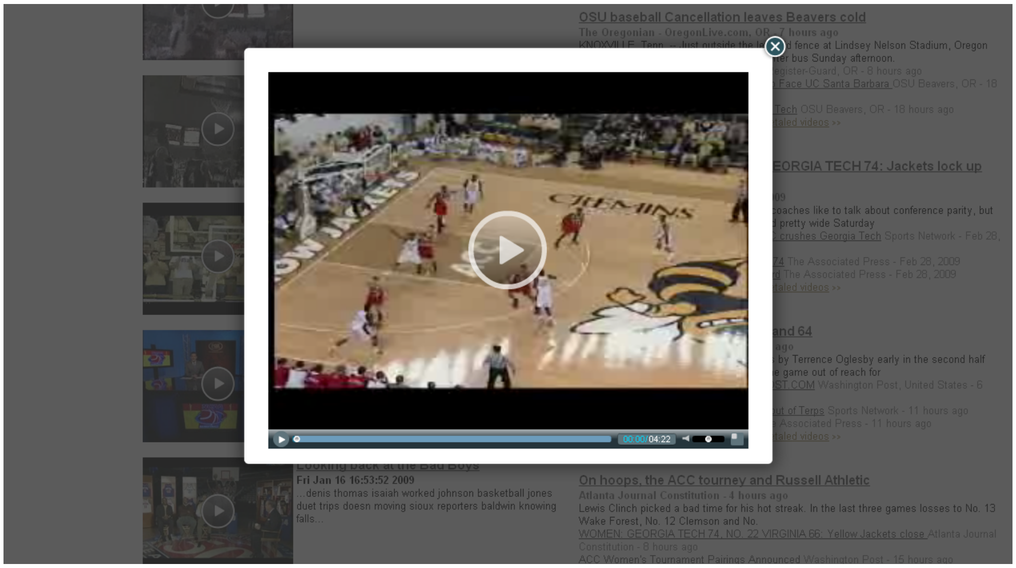
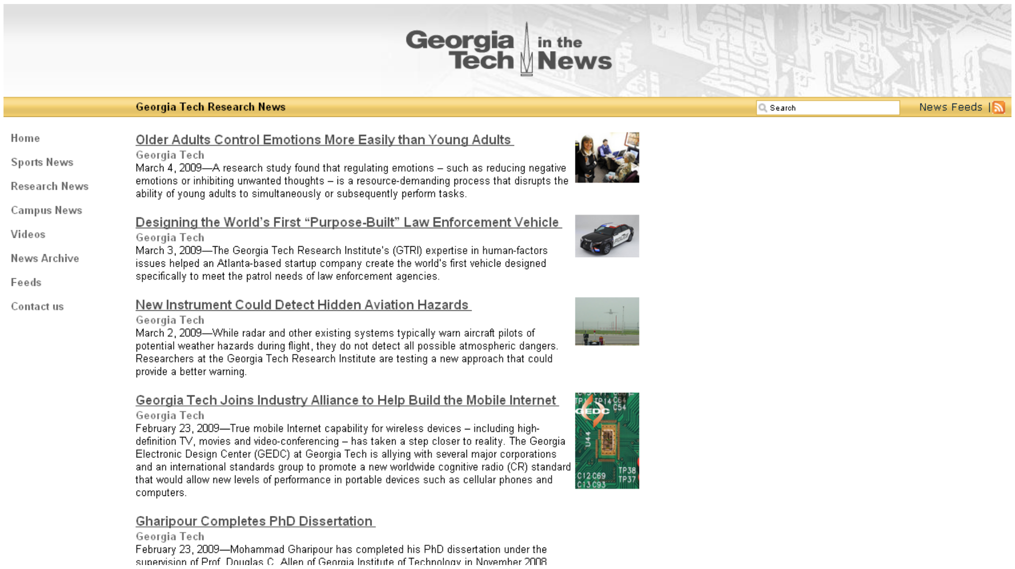

OSU baseball Cancellation leaves Beavers cold
The Oregonian - OregonLive.com, OR - 7 hours ago
d fence at Lindsey Nelson Stadium, Oregon
ter bus Sunday afternoon.
EORGIA TECH 74: Jackets lock up
coaches like to talk about conference parity, but
d pretty wide Saturday
crushes Georgia Tech Sports Network - Feb 28,
74 The Associated Press - Feb 28, 2009
rd The Associated Press - Feb 28, 2009
and 64
s by Terrence Oglesby early in the second half
e game out of reach for
00:00/04:22
Looking back at the Bad Boys
Fri Jan 16 16:53:52 2009
...denis thomas isaiah worked johnson basketball jones duet trips doesn moving sioux reporters baldwin knowing falls...
On hoops, the ACC tourney and Russell Athletic
Atlanta Journal Constitution - 4 hours ago
Lewis Clinch picked a bad time for his hot streak. In the last three games losses to No. 13 Wake Forest, No. 12 Clemson and No.
WOMEN: GEORGIA TECH 74, NO. 22 VIRGINIA 66: Yellow Jackets close Atlanta Journal Constitution - 8 hours ago
ACC Women's Tournament Pairings Announced Washington Post - 15 hours ago

Figure 14: Example of video transcoding



Georgia Tech Research News

Older Adults Control Emotions More Easily than Young Adults
Georgia Tech
March 4, 2009—A research study found that regulating emotions – such as reducing negative emotions or inhibiting unwanted thoughts – is a resource-demanding process that disrupts the ability of young adults to simultaneously or subsequently perform tasks.

Designing the World's First "Purpose-Built" Law Enforcement Vehicle
Georgia Tech
March 3, 2009—The Georgia Tech Research Institute's (GTRI) expertise in human-factors issues helped an Atlanta-based startup company create the world's first vehicle designed specifically to meet the patrol needs of law enforcement agencies.

New Instrument Could Detect Hidden Aviation Hazards
Georgia Tech
March 2, 2009—While radar and other existing systems typically warn aircraft pilots of potential weather hazards during flight, they do not detect all possible atmospheric dangers. Researchers at the Georgia Tech Research Institute are testing a new approach that could provide a better warning.

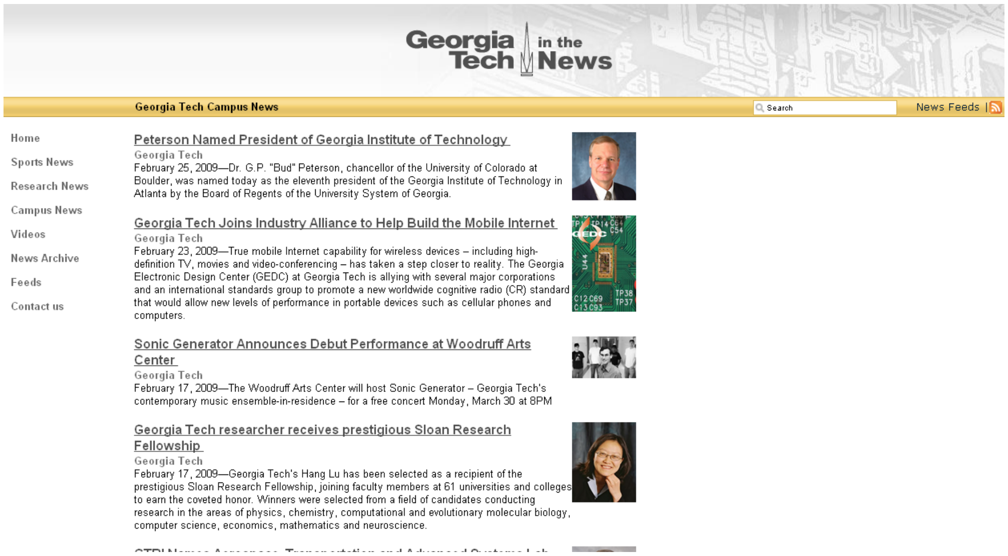
Georgia Tech Joins Industry Alliance to Help Build the Mobile Internet
Georgia Tech
February 23, 2009—True mobile Internet capability for wireless devices – including high-definition TV, movies and video-conferencing – has taken a step closer to reality. The Georgia Electronic Design Center (GEDC) at Georgia Tech is allying with several major corporations and an international standards group to promote a new worldwide cognitive radio (CR) standard that would allow new levels of performance in portable devices such as cellular phones and computers.

Gharipour Completes PhD Dissertation
Georgia Tech
February 23, 2009—Mohammad Gharipour has completed his PhD dissertation under the supervision of Prof. Douglas C. Allen of Georgia Institute of Technology in November 2008

Figure 15: Example of research news clusters related to Georgia Tech

5.1 Feature Comparison

Tables 2 and 3 show feature comparison charts of our proposed system with YouTube and Google News.



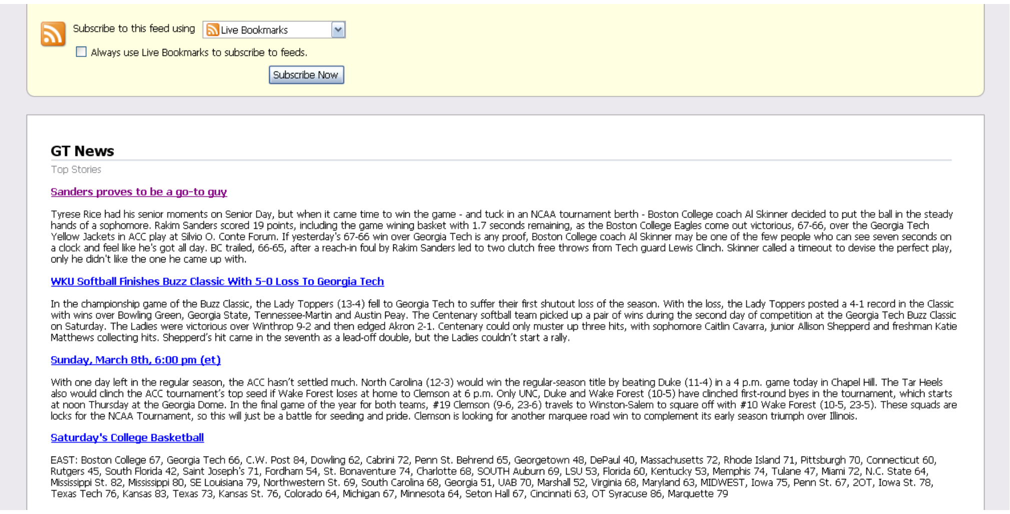
Peterson Named President of Georgia Institute of Technology
Georgia Tech
February 25, 2009—Dr. G.P. "Bud" Peterson, chancellor of the University of Colorado at Boulder, was named today as the eleventh president of the Georgia Institute of Technology in Atlanta by the Board of Regents of the University System of Georgia.

Georgia Tech Joins Industry Alliance to Help Build the Mobile Internet
Georgia Tech
February 23, 2009—True mobile Internet capability for wireless devices – including high-definition TV, movies and video-conferencing – has taken a step closer to reality. The Georgia Electronic Design Center (GEDC) at Georgia Tech is allying with several major corporations and an international standards group to promote a new worldwide cognitive radio (CR) standard that would allow new levels of performance in portable devices such as cellular phones and computers.

Sonic Generator Announces Debut Performance at Woodruff Arts Center
Georgia Tech
February 17, 2009—The Woodruff Arts Center will host Sonic Generator – Georgia Tech's contemporary music ensemble-in-residence – for a free concert Monday, March 30 at 8PM

Georgia Tech researcher receives prestigious Sloan Research Fellowship
Georgia Tech
February 17, 2009—Georgia Tech's Hang Lu has been selected as a recipient of the prestigious Sloan Research Fellowship, joining faculty members at 61 universities and colleges to earn the coveted honor. Winners were selected from a field of candidates conducting research in the areas of physics, chemistry, computational and evolutionary molecular biology, computer science, economics, mathematics and neuroscience.

Figure 16: Example of campus news clusters related to Georgia Tech


Subscribe to this feed using
Live Bookmarks
□ Always use Live Bookmarks to subscribe to feeds.
Subscribe Now

GT News
Top Stories

Sanders proves to be a go-to guy

Tyrese Rice had his senior moments on Senior Day, but when it came time to win the game - and tuck in an NCAA tournament berth - Boston College coach Al Skinner decided to put the ball in the steady hands of a sophomore. Rakim Sanders scored 19 points, including the game wining basket with 1.7 seconds remaining, as the Boston College Eagles come out victorious, 67-66, over the Georgia Tech Yellow Jackets in ACC play at Silvio O. Conte Forum. If yesterday's 67-66 win over Georgia Tech is any proof, Boston College coach Al Skinner may be one of the few people who can see seven seconds on a clock and feel like he's got all day. BC trailed, 66-65, after a reach-in foul by Rakim Sanders led to two clutch free throws from Tech guard Lewis Clinch. Skinner called a timeout to devise the perfect play, only he didn't like the one he came up with.

WKU Softball Finishes Buzz Classic With 5-0 Loss To Georgia Tech

In the championship game of the Buzz Classic, the Lady Toppers (13-4) fell to Georgia Tech to suffer their first shutout loss of the season. With the loss, the Lady Toppers posted a 4-1 record in the Classic with wins over Bowling Green, Georgia State, Tennessee-Martin and Austin Peay. The Centenary softball team picked up a pair of wins during the second day of competition at the Georgia Tech Buzz Classic on Saturday. The Ladies were victorious over Winthrop 9-2 and then edged Akron 2-1. Centenary could only muster up three hits, with sophomore Caitlin Cavarra, junior Allison Shepperd and freshman Katie Matthews collecting hits. Shepperd's hit came in the seventh as a lead-off double, but the Ladies couldn't start a rally.

Sunday, March 8th, 6:00 pm (et)

With one day left in the regular season, the ACC hasn't settled much. North Carolina (12-3) would win the regular-season title by beating Duke (11-4) in a 4 p.m. game today in Chapel Hill. The Tar Heels also would clinch the ACC tournament's top seed if Wake Forest loses at home to Clemson at 6 p.m. Only UNC, Duke and Wake Forest (10-5) have clinched first-round byes in the tournament, which starts at noon Thursday at the Georgia Dome. In the final game of the year for both teams, #19 Clemson (9-6, 23-6) travels to Winston-Salem to square off with #10 Wake Forest (10-5, 23-5). These squads are locks for the NCAA Tournament, so this will just be a battle for seeding and pride. Clemson is looking for another marquee road win to complement its early season triumph over Illinois.

Saturday's College Basketball

EAST: Boston College 67, Georgia Tech 66, C.W. Post 84, Dowling 62, Cabrini 72, Penn St. Behrend 65, Georgetown 48, DePaul 40, Massachusetts 72, Rhode Island 71, Pittsburgh 70, Connecticut 60, Rutgers 45, South Florida 42, Saint Joseph's 71, Fordham 54, St. Bonaventure 74, Charlotte 68, SOUTH Auburn 69, LSU 53, Florida 60, Kentucky 53, Memphis 74, Tulane 47, Miami 72, N.C. State 64, Mississippi St. 82, Mississippi 80, SE Louisiana 79, Northwestern St. 69, South Carolina 68, Georgia 51, UAB 70, Marshall 52, Virginia 68, Maryland 63, MIDWEST, Iowa 75, Penn St. 67, 2OT, Iowa St. 78, Texas Tech 76, Kansas 83, Texas 73, Kansas St. 76, Colorado 64, Michigan 67, Minnesota 64, Seton Hall 67, Cincinnati 63, OT Syracuse 86, Marquette 79

Figure 17: Example of custom RSS feeds

Table 2: Feature Comparison with YouTube

Feature	YouTube	Our System
Video search	Yes	Yes
Automated video clustering	Yes	Yes
Dynamic updates for videos	Yes	Yes
Video context information	Yes	Yes
RSS feeds	Yes	Yes

Table 3: Feature Comparison with Google News

Feature	Google News	Our System
News clustering by topic	Yes	Yes
News clustering by category	Yes	Yes
News summaries	No	Yes
Related news images	Yes	Yes
Related videos	No	Yes
News text search	Yes	Yes
News videos search	No	Yes
RSS feeds	Yes	Yes

CHAPTER VI

APPLICATIONS OF VIDEO CONTEXT

Video search engines are more complex than text search engines due to the challenges in searching videos based on metadata which may not provide comprehensive information about the videos. A lot of research has gone into the text search engines and they have been perfected over the years. Text search is useful when the user is looking for direct answers to some questions which are available in textual form in webpages or searching for some specific textual content such as research papers, blogs, news articles, etc. However, for some applications video search engines are more suitable as compared to the text search engines. Users find it easier to assimilate short videos which provide relevant information than browsing a number of text articles. Videos are useful to answer user queries which are not directed for some specific information, but intended to get some broad overview on some topic. For example, a query to find: by how many runs did India win a cricket match, can be answered better by text search engines, as they search for specific keywords from the user query in the webpages and provide the search results. However, a query: highlights of cricket match, which is more generic in nature and not intended for some specific information or answers to some specific questions can be better answered by video search engines. The video search engines can return as the search results some short clips showing the highlights of the cricket match.

The technologies for context based video search described in Chapter IV have a number of applications as described below.

6.1 News Video Search

As described in Chapter IV, the context based video search technologies are very useful for searching for dynamic content such as news videos. Thousands of news videos are uploaded on the news video broadcasting websites; however, due to the lack of the metadata and the tags associated with videos, searching for such videos becomes difficult. There is no such search engine available currently that clusters dynamic content such as news videos from a number of sources and makes it searchable. Moreover, searching videos based on keywords from the metadata and user tags may not provide relevant search results, as the keywords may not of effectively linked to the video content. Searching news videos based on the low-level visual features extracted from the key-frames of the videos is also not effective, as many news videos may have similar visuals. For example, news videos on a cricket series from a particular news channel will have similar visual content for the entire cricket series, due to similar setup of the news studio and similar cricket grounds. Context based video search technologies presented in this thesis provide the most effective and efficient solutions for news video search for the following reasons.

6.1.1 News Video Repository Generation

The framework described in chapter IV can be used to generate news video repositories in an automated manner by companies, institutes and organizations for news videos relevant to them. In addition to that the framework can also be useful for generating repositories for news videos related to some specific area such as football news videos, etc. The generation of news repositories is guided by some specific contexts which are generated from clusters of news text articles. This approach overcomes the problem of searching videos based on metadata and user tags which do not provide much details about the video.

6.1.2 Context Based Search

Searching videos based on the context of the news story provides more relevant results as compared to search based on the low-level features extracted from the videos or video content in the form of automatic speech recognition text.

6.1.3 Dynamic Ranking for Videos

The ranking of the search results is dynamic in nature. The VideoRank described in chapter IV, used for ranking the video search results depends on the relevance of the videos to some context. As newer and more relevant news videos become available, the VideoRank of a video for some context keeps changing, reflecting the relevance of the video to the news story.

6.1.4 Linking Videos to Relevant News Articles

The framework described in chapter IV not just allows building video repositories and searching videos based on context, but also links the news videos to the relevant news articles and news clusters. This novel approach is useful to generate webpages which have both the news stories and the relevant videos at the same place. Moreover, searching for a specific news story returns both the news articles and the related videos in the search results.

6.2 Search for Video Lectures

Many institutes and universities provide their video lectures online along with the lecture slides. Using the technologies for context based video search, effective search engines for video lectures can be built. These technologies are useful for video lecture search for the following reasons:

6.2.1 Clustering Video Lectures

The online video lecture websites have lectures on many subjects and courses which are offered over the years. These videos can be clustered based on the subject. For example, all video lectures on image processing courses, talks and workshops can be clustered which can make it easier for the users to find all relevant content on image processing at one place. The clustering process can be guided by the contexts generated from textual content of the lecture slides and notes.

6.2.2 Linking Video Lectures to Relevant Documents

The framework described in chapter IV can be used to link the video lectures to the relevant documents, lecture slides, notes, etc. For example, a department of some university can use the framework to cluster all the videos and documents on their website and generate webpages in an automated manner which have the video lectures clustered by subject and the related documents in one place. This can be an effective resource for students who, instead of browsing many webpages and searching of content on some specific subject, can find all the relevant material clustered in one webpage.

6.3 Video Advertisements

The technologies for context based video search can also be used for video advertisements. For example, to design an automated system that selects the most relevant video advertisements that are responsive to a query at a particular moment in time, and embeds them inside videos. Chapter VI describes a detailed approach and framework for using the video contexts for advertisements.

CHAPTER VII

FUTURE WORK

The technologies described in Chapter IV are not just useful for video search, but they can also be applied to a number of other applications. In this chapter we describe one such application of video context that is used to generate video advertisements. The objective of this application is to design an automated system that selects the most relevant video advertisements that are responsive to a query at a particular moment in time, and embeds them inside videos. A context-sensitive video search and advertisement selection scheme is used, wherein the context is generated in an automated manner.

7.1 Introduction

Internet advertising has seen tremendous growth in the past few years with online advertising spending estimated to be over $25 billion in US alone and $45 billion globally in 2008. The major share of the web advertising market today consists of textual ads which are placed either in the web pages or on the search result pages from the web search engines. The current approaches to text based online advertising are described below:

7.1.1 Contextual Advertising

This is a form of targeted advertising where the textual ads are placed on websites or email messages that have similar content. The advertisements are selected and served by the advertising networks (Google, Yahoo, MSN, AOL) which have automated systems of selecting the advertisements displayed to the user based on the content. These ads are believed to have a greater chance of attracting a user, because they tend to have similar content as the websites on which they are placed. The effectiveness of contextual advertisements depends on ad-selection techniques adopted by the ad-networks.

7.1.2 Search advertising

This is a method of placing advertisements on search result pages of web search engines based on the keywords from the user query. Search advertising is provided by search engines like Google and Yahoo that deliver ads on the basis of search keywords. Search engines conduct running auctions to sell ads according to bids received for keywords.

7.1.3 Inline advertising

This is another form of targeted advertising where the ads are delivered inside the webpage content. Unlike contextual advertisements where ads are placed at pre-defined portions of a webpage, the inline advertisements are embedded in the text of a webpage. The advertising networks associate relevant ads with certain keywords of the webpage text and highlight those keywords. The inline advertisements are shown when a user moves the mouse over these highlighted keywords.

7.2 Using Video Context for Advertisements

Since forms of online advertising are driven by keywords, textual ads comprise the major share of online advertisements. Video is another medium of advertising which is becoming poplar for online advertisements, as it is much more attractive and can grab users' attention instantly. Unlike textual ads where the keywords from the search query or the content of a webpage are used to select the advertisements, for in-video advertising, selecting the most relevant video advertisements that match the context of the video remains a challenging task. Commercial video search engines use text annotations and captions for video search and advertisement selection. However keywords are not sufficient for effectively selecting the most relevant advertisements for videos.

An outcome of this approach is that for a particular query, the same search results are produced, irrespective of the users context of search. However, the context underlying the search for each user may be entirely different. For example, a user in California who searches a term like pizza is more interested in pizza advertisements from California rather than other places. The current search engines provide search results which are same for all the users, hence the same advertisements are displaced for every user irrespective of the context of the users query. A system which does a context sensitive ranking of search results can provide much more meaningful search results and related advertisements to a user. The effectiveness of an online advertisement is usually defined from the advertisers' perspective, and measured by the performance of a given advertisement (e.g. the number of clicks). Thus advertisements that are contextually relevant to a user's query are more likely to attract attention and prove to be more effective from the advertisers perspective.

Current contextual advertising approaches work well for inserting textual ads into static web pages where the webpage content is analyzed in advance and keywords are extracted from the content to associate advertisements with the webpage. However for web pages having dynamic content like news and videos, analyzing the content on the fly is computationally intensive and introduces significant latencies.

We propose a novel context based video clustering, search and advertising approach that attempts to make the generation of automated real time video repositories efficient, and also tries to make the process of video search and advertisement selection more meaningful. The system can be deployed as a software as a solution by organizations and institutes to build online video repositories and provide effective video search and in-video advertising capabilities. For example, a sports news agency that produces both text and video news can use this system to organize their videos by linking them to the most relevant text articles and provide contextually relevant advertisements embedded inside the videos.

7.3 Framework

7.3.1 Video Clustering

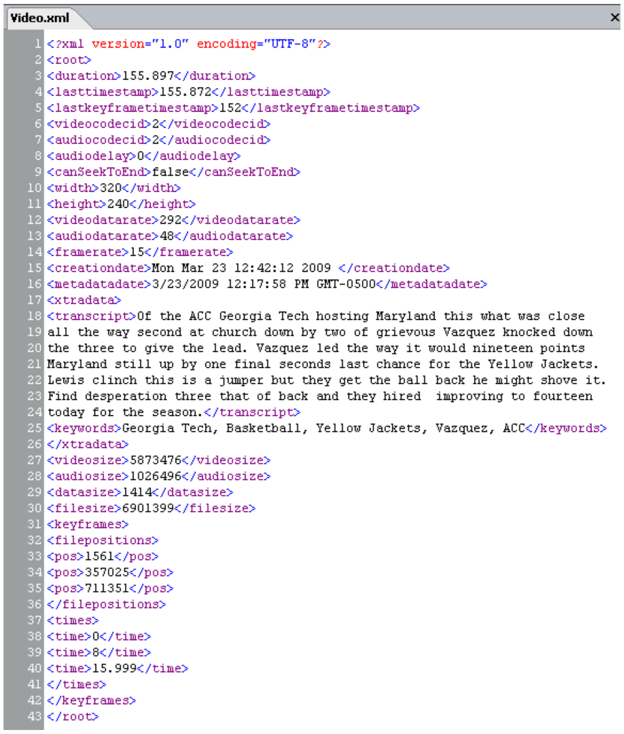
To build video repositories from the videos available on the internet or the intranet in an automated manner our system first clusters text articles based on topic or category and uses the contexts generated from similar text articles to cluster related videos. The idea is that the video archives available on the internet or the intranet may be unorganized and may not have enough textual information attached to them to derive a complete context of the video. Similarly, the text based articles may not be directly linked to the video archives. For example a sports agency that produces both text and video news may have huge text and video news archives which are unorganized and there is no way to link the text articles to the relevant videos. Our system tries to bridge the gap by providing a mapping of the text articles to the related videos.

7.3.2 Context generation

Our system uses a query expansion technique to enhance contextual information of the query which can then be used to retrieve the most relevant videos and also map the contextually relevant advertisements to the videos. The idea is to analyze the query terms and find other similar terms or keywords which have a high correlation with the query terms. The context generation module generates a word cloud from the query which has words that are highly correlated to the query. For example, a user query like "ACC" provides very little information about the user's context of search. Assuming that the system is deployed for a sports news agency, the context generation module will generated a word cloud from the query having words like Atlantic Coast Conference, ACC basketball, sports, tournament, championship, etc.

To generate the context the system analyses the text articles and word clouds of videos in the database and chooses the words which are highly correlated to the search query. So in the above example, its possible that the system had some text article which had words like ACC, basketball, etc. therefore the system was able to attach such correlated words to the query to form a word cloud.

The system wraps each video with a rich layer of metadata information that includes the ASR text and a word cloud generated using the keywords from the ASR text. Figure 18 shows an example of the video metadata information generated for a video in XML format.



```xml
<?xml version="1.0" encoding="UTF-8"?>
<root>
<duration>155.897</duration>
<lasttimestamp>155.872</lasttimestamp>
<lastkeyframetimestamp>152</lastkeyframetimestamp>
<videocodecid>2</videocodecid>
<audiocodecid>2</audiocodecid>
<audiodelay>0</audiodelay>
<canSeekToEnd>false</canSeekToEnd>
<width>320</width>
<height>240</height>
<videodatarate>292</videodatarate>
<audiodatarate>48</audiodatarate>
<framerate>15</framerate>
<creationdate>Mon Mar 23 12:42:12 2009 </creationdate>
<metadatadate>3/23/2009 12:17:58 PM GMT-0500</metadatadate>
<xtradata>
<transcript>Of the ACC Georgia Tech hosting Maryland this what was close
all the way second at church down by two of grievous Vazquez knocked down
the three to give the lead. Vazquez led the way it would nineteen points
Maryland still up by one final seconds last chance for the Yellow Jackets.
Lewis clinch this is a jumper but they get the ball back he might shove it.
Find desperation three that of back and they hired  improving to fourteen
today for the season.</transcript>
<keywords>Georgia Tech, Basketball, Yellow Jackets, Vazquez, ACC</keywords>
</xtradata>
<videosize>5873476</videosize>
<audiosize>1026496</audiosize>
<datasize>1414</datasize>
<filesize>6901399</filesize>
<keyframes>
<filepositions>
<pos>1561</pos>
<pos>357025</pos>
<pos>711351</pos>
</filepositions>
<times>
<time>0</time>
<time>8</time>
<time>15.999</time>
</times>
</keyframes>
</root>
```

Figure 18: An example of the rich meta-data information generated for a video

7.3.3 Video Advertisement Insertion

The word cloud from the expanded user query is used to search for videos and also find the relevant advertisements for embedding into the videos. Our system assigns relevance weights to the videos with respect to the query context and calculates the

VideoRankTM, which indicates the likelihood of the presence of a context in a video. The advertisement insertion module generates a set of candidate advertisements that are relevant to the query context and embeds them into the retrieved videos.

7.4 System Architecture

Figure 19 illustrates the system overview. The user query first goes to the context generation module. The context generation module creates a word cloud around the query which is then sent to the index servers. Index servers have indexes of video metadata including the video ASR text and the video context. The query is then sent to the context servicing and VideoRank agent which ranks the videos based on the query context. The content delivery servers retrieve the videos and generate the content describing the video search result. Also the generated context of the query is sent to the advertisements keyword index server, which selects the most relevant video advertisements based on the context. The keywords of advertisements are extracted from the advertisement titles, text and categories provided by advertisers. The selected advertisements are then embedded into the videos by the video ads insertion module. The video search results with embedded video advertisements are then returned to the user.

7.5 Sample Results

Figure 20 illustrates the video search and advertisement selection and insertion process for the query ACM.

The user query is expanded to derive the context and a word cloud from the query is generated. To retrieve the videos the system finds intersections between the word cloud of the query and the video word clouds which are in the form of an XML based metadata layer over the video as shown in figure 18.The advertisements are then selected based on the query word cloud and embedded into the videos.

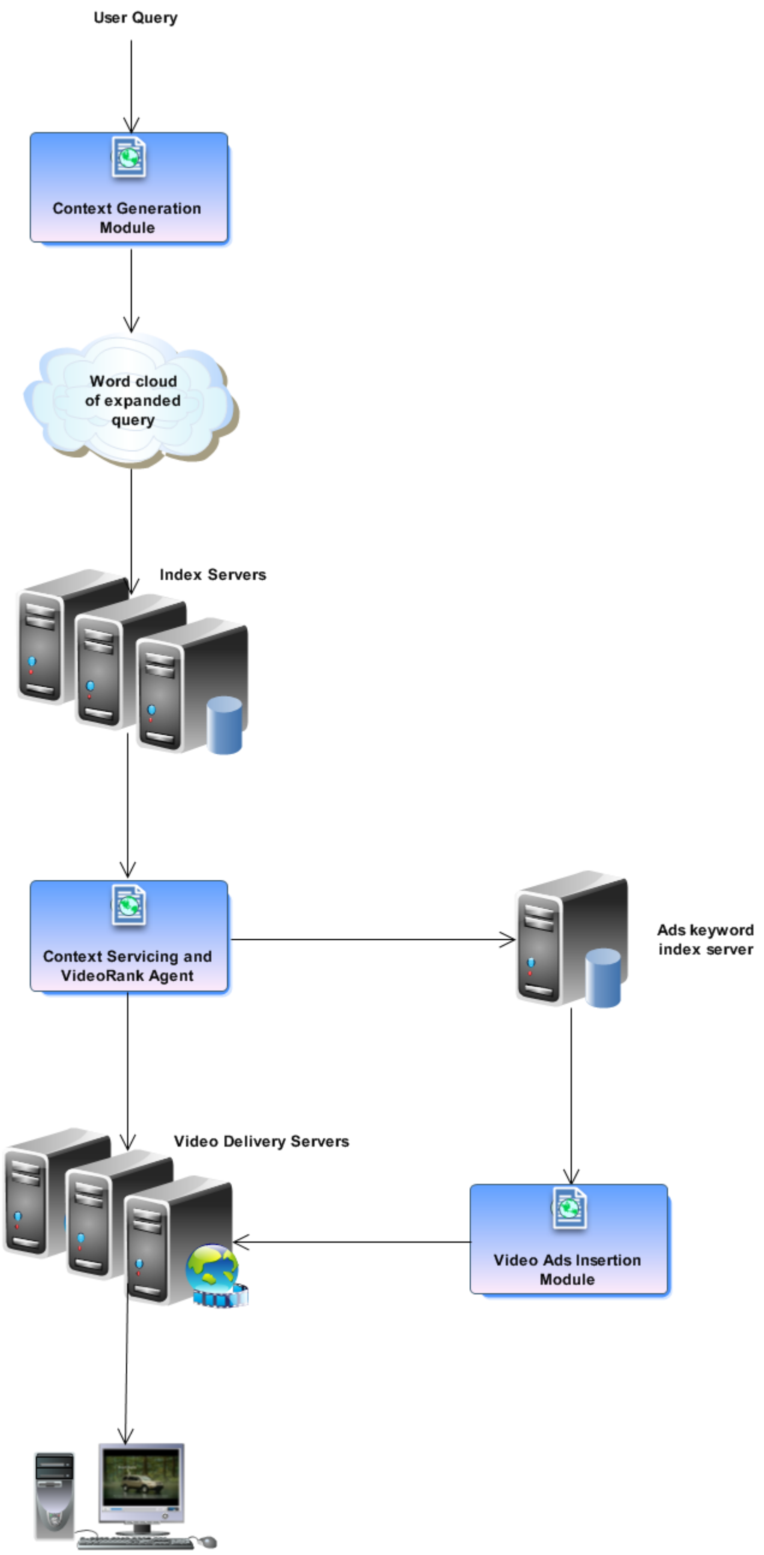

User Query
Context Generation Module
Word cloud of expanded query
Index Servers
Context Servicing and VideoRank Agent
Ads keyword index server
Video Delivery Servers
Video Ads Insertion Module

Figure 19: Context based video advertisement system architecture

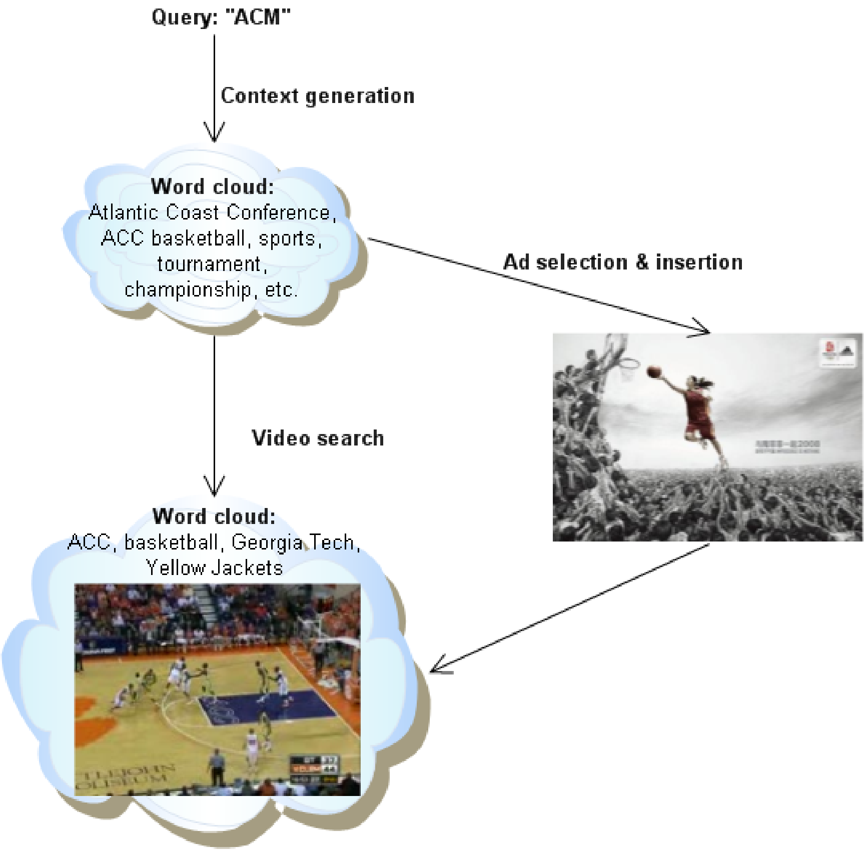


Figure 20: An example of video search and ad selection based on the query ACM

CHAPTER VIII

CONCLUSION

This thesis proposes technologies for context based video search. A framework for building video repositories in an automated manner guided by contexts that are obtained from clusters of text articles is described. A context sensitive video ranking is proposed which is used to rank the video search results. The context based video search approach has applications in searching dynamic content such as news videos, video lectures, and video advertisements. Context-based video search also has applications in areas such as digital video broadcasts, video on demand, and video surveillance. Most applications require the ability to search videos based on the semantic and contextual information available. A limitation of content based video search approaches is that it is difficult to relate the low level features with semantics. The commercial video search engines available today depend on the users to upload videos manually and the search is again dependent on the captions and annotations provided by the user. Thus they are not able to keep up with the rapid rate at with which new videos are being uploaded by various video broadcasts websites as there is a complete lack of automation. There is an increasing demand for online video broadcast services, driven primarily by the ease with which users can access videos through mobile phones, PDAs and other hand held devices. Thus, a real time and automated video clustering and search system which not only provides the users the most relevant videos available at a particular moment but also the related contexts of the videos, summarized from a number of different sources, will become indispensable for users in future.

REFERENCES

[1] D. Radev, J. Otterbacher, A. Winkel, and S. Blair-Goldensohn. NewsInEssence: Summarizing online news topics, Communications of the ACM, Volume 48 , Issue 10, October 2005.

[2] D. K. Evans, J. L. Klavans, and K. R. McKeown. Columbia newsblaster: Multilingual news summarization on the web. Human Language Technology (HLT), Boston, MA, May, 2004.

[3] V. Thapar, A.A. Mohamed, and S. Rajasekaran. A consensus text summarizer based on meta-search algorithms. IEEE International Symposium on Signal Processing and Information Technology, Vancouver, British Columbia, Australia, 2006.

[4] H. Geng, P. Zhao, E. Chen, and Q. Cai. A novel automatic text summarization study based on term co-occurrence. 5th IEEE International Conference on Cognitive Informatics, Beijing, 2006.

[5] O. Sornil and K. Gree-ut. An automatic text summarization approach using content-based and graph-based characteristics. IEEE Conference on Cybernetics and Intelligent Systems, Bangkok, 2006.

[6] L. Yu, J. Ma, F. Ren, and S. Kuroiwa. Automatic text summarization based on lexical chains and structural features, Proceedings of the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing - Volume 02, 2007.

[7] K. Kaikhah. Automatic text summarization with neural networks. Second IEEE International Conference on Intelligent Systems, June 2004.

[8] M. T. Maybury and A. E. Merlino. Multimedia summaries of broadcast news, Proceedings on Intelligent Information Systems, Grand Bahama Island, Bahamas, pp, 442-449, 1997.

[9] A. Chongsuntornsri and O. Sornil. An automatic thai text summarization using topic sensitive PageRank, International Symposium on Communications and Information Technologies, Bangkok, pp. 574-552, 2006.

[10] C. Li and S. Wang. Study of automatic text summarization based on natural language understand- ing, IEEE International Conference on Industrial Informatics, Singapore, pp. 712-714, 2006.

[11] S. Y. Neo, J. Zhao, M. Y. Kan, and T. S. Chua, Video search using high-level features: Exploiting query-matching and confidence-based weighting, CIVR 2006, Tempe, AZ, pp. 143-152, July 2006.

[12] H. Yang, L. Chaisorn, Y. Zhao, S.-Y. Neo, and T. S. Chua, VideoQA: Question answering on news video ACM Multimedia 2003, Berkeley, CA, pp. 632-641, Nov 2003.

[13] J. McCarley and M. Franz. Influence of speech recognition errors on topic detection. SIGIR 2000, 342-344, New York, 2000.

[14] J. Allan, R. Papka, and V. Lavrenko, On-line new event detection and tracking SIGIR 1998, Melbourne, Australia, 37-45, 1998.

[15] W. H. Hsu, L. Kennedy, S. F. Chang, M. Franz, and J. Smith, Columbia-IBM news video story segmentation in TRECVID 2004, Columbia ADVENT Technical Report, 2005.

[16] A. P. Natsev, M. R. Naphade, and J. R. Smith, Semantic representation, search and mining of multimedia content, ACM Int. Conf. on Knowledge Discovery and Datamining (SIGKDD), Seattle, WA, 2004.

[17] S.Y. Neo, J. Zhao, M.-Y. Kan, and T.S. Chua, Video search using high level features: Exploiting query matching and confidence-based weighting, Int. Conf. on Image and Video Search (CIVR), Tempe, AZ, 2006.

[18] M. Campbell, S. Ebadollahi, D. Joshi, M. Naphade, A. P. Natsev, J. Seidl, J. R. Smith, K. Scheinberg, J. Tesic, L. Xie, and A. Haubold, IBM research TRECVID-2006 video search system, TRECVID, Gaithersburg, MD, 2006.

[19] C. G. M. Snoek, B. Huurnink, L. Hollink, M. de Rijke, G. Schreiber, and M. Worring, Adding semantics to detectors for video search, IEEE Trans. Multimedia, 9(5), 975-986, 2007

[20] S. F. Chang, W. Hsu, W. Jiang, L. Kennedy, D. Xu, A. Yanagawa, and E. Zavesky, Columbia university TRECVID-2006 video search and high-level feature extraction, in TRECVID, Gaithersburg, MD, 2006.

[21] M. Naphade, J. R. Smith, J. Tesic, S.-F. Chang, W. Hsu, L. Kennedy, A. Hauptmann, and J. Curtis, Large-scale concept ontology for multimedia, IEEE Multimedia, 13(3), 86-91, 2006

[22] C.G.M. Snoek, J.C. van Gemert, J. M. Geusebroek, B. Huurnink, D.C. Koelma, G.P. Nguyen, O. De Rooij, F. J. Seinstra., A. W. M. Smeulders, C. J. Veenman., and

M. Worring, The MediaMill TRECVID 2005 semantic video search engine, TRECVID Workshop, NIST, Gaithersburg, MD, Nov 2005.

[23] A. Hauptmann., M. Christel, R. Concescu, J. Gao, Q. Jin, W. H. Lin, J. Y. Pan, S. M. Stevens, R. Yan, J. Yang, and Y. Zhang, CMU Informedias TRECVID 2005 skirmishes, TRECVID Workshop, NIST, Gaithersburg, MD, Nov 2005.

[24] H. Yang, T.-S. Chua, S. Wang, and C.-K. Koh, Structured use of external knowledge for event-based open-domain question answering, SIGIR 2003, Toronto, Ontario, Canada, pp. 33-40, July 2003.

[25] M. Petkovic, Content-based Video Retrieval, VII. Conference on Extending Database Technology (EDBT), Ph.D. Workshop, Konstanz, Germany, March 2000.

[26] A. Hampapur, A. Gupta, B. Horowitz, C-F. Shu, C. Fuller, J. Bach, M. Gorkani, and R. Jain, Virage video engine, SPIE, 3022, 188-198, 1997.

[27] D. Ponceleon, S. Srinivasan, A. Amir, D. Petkovic, and D. Diklic, Key to effective video search: effective cataloging and browsing, ACM Multimedia98, Bristol, U.K., pp. 99-107, 1998.

[28] S.F. Chang, W. Chen, H. Meng, H. Sundaram, and D. Zhong, A fully automated content based video search engine supporting spatio-temporal queries, IEEE Transaction Circ. Syst. Video Technol., 8(5), 302-615, Sept., 1998.

[29] K. Wan, Exploiting story-level context to improve video search, IEEE International Conference on Multimedia and Expo, Hannover, Germany, pp. 298-292, 2008.

[30] S.Y. Neo, Y. Ran, H.K. Goh, Y. Zheng, T.S. Chua, and J. Li, The use of topic evolution to help users browse and find answers in news video corpus, Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany, pp. 198-207, 2008.

[31] S. Brin and L. Page, The anatomy of a large-scale hypertextual web search engine, Computer Networks and ISDN Systems, Volume 30 , Issue 1-7, April 1998.

VITA

Arshdeep Bahga was born in Chandigarh, India on December 24, 1983. He holds a Bachelor of Engineering degree in Electronics and Electrical Communication from Punjab Engineering College, Chandigarh. He has worked as a Software Engineer in Electronics for Imaging, Inc in Bangalore, India for two years before he joined Georgia Tech for MS in Fall 2008. His areas of interest include Digital Signal and Image Processing. He has worked on research projects in the area of image and video retrieval for two years.