International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 04 | Apr 2019 www.irjet.net p-ISSN: 2395-0072

Segmenting, Multimedia Summarizing and Query Based Retrieval of News Story from News Broadcast

Mr. N. V. Bhalerao[1], Dr. Mrs. S. S. Apte[2], Dr. Mrs. A. R. Kulkarni[3]

[1] ME Student, Dept. of Computer Science and Engineering, WIT, Maharashtra, India
[2] Professors, Dept. of Computer Science and Engineering, WIT, Maharashtra, India
[3] Professors, Dept. of Computer Science and Engineering, WIT, Maharashtra, India

---

Abstract- *This project builds a system to summarize and retrieve the broadcast news at multimedia level. This project combines anchor person based story boundary detection and text summarization system to build multimedia news summary and news extraction system. Broadcast news are captured both in video/audio with accompanying transcript in text format. Summarized individual news story video clip according to textual summary can be retrieved by user query. Lexical chain text summarization technique is used to summarize individual news story transcript. The summary is in multimedia format including video, audio, and text.*

***Key Words*: Multimedia Summarization, Retrieval System, Story Boundary Detection, Text Summarization, Lexical Chain, News Broadcast.**

1. INTRODUCTION

Broadcast news video has been playing all the time more vital role in our everyday life. As a kind of primary multimedia resources, broadcast news videos are regularly accessed by a large number of people all over the world. The broadcast news video has several distinct characteristics that are quite different with other types of videos. For instance, a story unit always accompanies with one to several descriptive caption texts. Anchor persons usually appear at the beginning of a story unit. Thus, the broadcast news video can be regarded as a kind of semi-structured multimedia data that contains informative clues for parsing itself into semantic story units.

Broadcast news videos are regularly accessed by a large number of people all over the world. However, despite its popularity, finding video of a particular person's interest is still by no means easy for several reasons. First, broadcast news video archives contain a large number of historical videos. Facing the overwhelming amounts of videos, people are always difficult to locate the videos they are interested in, and thus "lost in TV program space". Second, the latest broadcast news videos are delivered daily at fixed time, people sometimes may not convenience to watch them due to time conflict. In addition, there are cases that two videos a person both interested in are delivered as the same time. Third, instead of watching the whole video, sometimes what people want to watch is only a particular segment of interest.

To efficiently manipulate and manage the increasing broadcast news videos, a key technique is by segmenting the whole video into meaningful and relatively independent video clips each depicting only one story, i.e., story units.

The aim is to build broadcast news summarization system. It takes YouTube news broadcast, analysis the content to identify news stories. Content of each story is summarized and important keywords are extracted. This data is to be stored in a database, and news Retrieval system is to be implemented which let users to explore for any part of news in the database. Such a system have improvement over other search engines, as we use summarization methods to focus the most important information to the user, facilitating him/her to find the story he/she is eyeing for in a smaller time, when compared to an ordinary text based search engine.

1.1 Aim of Project

- Story segmentation – in order to identify story boundaries from 30 minutes worth of video and transcript (subtitles). (E.g.to find that there are 10 individual pieces of news in a RajyaSabha TV news bulletin at night). A technique suitable to solve this problem is story segmentation using anchor person. It identifies the current story and detects when/where the story changes. It helps to detect the story boundaries and identifies story frames. There are known methods to carry out video segmentation but all of them suffer from

certain limitations, and this problem has been proved to be difficult to solve. We try to implement story segmentation with the use of anchor person frame to segment the video into news stories. Anchor person usually appear at the beginning of every news story in news bulletin.

- Story summarization – provides the user with a short summary of the story both in text and video. This lets the user to choose whether a video excerpt returned by the search engine is related to the subject he/she is searching for. We apply a text summarization technique called 'Lexical Chain' to summarize the text. Video summary of story is generated using summarized text.
- Keyword based news retrieval system-apart from a summary of the original story, we could also try to identify important keywords such as "Delhi", "Prime Minister" or "blast" from news transcript for each story segment. Such keywords, if recognized correctly, would contribute incredible amount to specify what the topic of the story is about. Based on frequency of occurrences of matching keywords user can select stories of interest.

Video analysis techniques are used to story segmentation, extracting prominent feature from video and summarizing the story video as per text summary. We have used natural language processing technique and tool (WordNet –lexical database) to extract important keyword and summarize the text.

2. LITERATURE RIVIEW

Mark T. Maybury and Andrew E. Merlino has implemented and extracted "Summaries for Broadcast News". They have used algorithm for summarization, key phrase extraction and story segmentation, and key frame extraction generate the summarized video.

Marcus J. Pickering, Stefan M. Rüger has focused "video search engine using dual-media segmentation". They implemented an algorithm which uses the audio track for identifying meaningful scene breaks. This work is related to web-based video search engine that is implemented using broadcast news, and the main part of implementation is story boundary detection.

Kuan-Yu Chen, Shih-Hung Liu has implemented "Extractive Broadcast News Summarization Leveraging Recurrent Neural Network Language Modeling Techniques" their work in this paper mainly focused on use of recurrent neural network language modeling (RNNLM) framework for extraction as well as summarization of broadcast news.

Mark T. Maybury mainly focused on "Discourse Cues for Broadcast News Segmentation" they describe analysis of a broadcast News corpus, and focused on information extraction techniques, and finally its computational implementation and evaluation in the Broadcast News Navigator (BNN) for achieving browsing, retrieval, and summarization of news video.

Warren Greiff, Alex Morgan, Randall Fish, Marc Richards, Amlan Kundu presented "Fine-Grained Hidden Markov Modeling for Broadcast-News Story Segmentation". The News broadcasts are divided into story segments by using Hidden Markov Model. Model topology and the textual features used together with the non-parametric approximation techniques for obtaining estimates for both transition and observation probabilities. Visualization approaches developed for the examination of system performance.

Kathleen McKeown and Dragomir R. Radev have established a model for "Generating Summaries of Multiple News Articles". They have offered Natural Language system for summarization of a sequence of News articles on the same event.

Hemant Misra, Frank Hopfgartner, Anuj Goyal, P. Punitha, and Joemon M. Jose are focused on "TV News Story Segmentation based on Semantic Coherence and Content Similarity". They have assessed two methodologies, one using video stream and the other using close-caption text stream, for segmenting TV news into stories. The segmentation of the video stream into stories is achieved by detecting anchor person shots and the text stream is segmented into stories using a Latent Dirichlet Allocation (LDA) based approach.

Regina Barzilay,Michael Elhadad have explored one method to summarize original text by using the model of the topic progression in the text resulting from lexical chain. They offered new algorithm to compute lexical chain in the text.

Zechao Li, Jinhui Tang ,Xueming Wang, Jing Liu , Hanqing Lu have developed new method of multimedia news summarization for searching results on the

Internet, which finds the essential topics among query-related news information and threads the news events within each topic to generate a query-related brief summary. They used HLDA to topic structure from query associated news document. And time influenced maximum spanning tree algorithm is suggested to form condensed summary of parent topic.

Peter Bell, Catherine Lai, Clare Llewellyn, Alexandra Birch, Mark Sinclair defined an end-to-end system for processing and browsing audio news data. Their fully automated system carries together recent research on audio scene analysis, speech recognition, and summarisation, named entity detection, geo location, and machine translation.

Jia-Yu Pan, Hyungjeong Yang, and Christos Faloutsos recommended multi-modal story-oriented video summarization (MMSS). MMSS discovers association between information of different modalities which gives expressive story-oriented news video summaries. MMSS can also be applied for video retrieval.

Ichiro Ide, Ye Zhang, Ryunosuke Tanishige, Keisuke Doman,Yasutomo Kawanishi, Daisuke Deguchi, and Hiroshi Murase put forward a method for summarizing a sequence of news videos considering the steadiness of both auditory and visual contents. The suggested technique first selects key-sentences from the auditory contents (Closed Caption) of each news story in the sequence, and then picks a shot within the news story whose "Visual Concepts" identified from the visual contents are the most consistent with the key-phrase. Finally, the audio segment matching to each key-phrase is coincided onto the selected shot, and then concatenated to generate a summarized video.

Mi-mi LU, Lei XIE, Zhong-hua FU, Dong-mei JIANG and Yan-ning ZHANG studied how to assimilate multi-modal features for story boundary detection in broadcast news. The uncovering problem is expressed as a classification task, i.e., classifying each candidate into boundary/non-boundary based on a set of features. They used a varied collection of features from text, audio and video modalities: lexical features capturing the semantic shifts of news topics and audio/video features reflecting the editorial rules of broadcast news.

Bailan Feng, Zhineng Chen, Rong Zheng, Bo Xu proposed a new unified video structure parsing method, named multiple style exploration-based news story segmentation (MSE-NSS), to segment broadcast news videos into semantic story units. In MSE-NSS, they first explore the suitable methods to explore various kinds of style information intrinsic in broadcast news videos, comprising temporal style inferred from caption texts, boundary style signified by a affluence of multimodal visual–audio features, and structural style known as the spanning duration of story units. The task of story unit segmentation is accomplished through the following three steps: temporal style-based pre-location, boundary style-based description, and boundary-structural style-based segmentation. Parallel to this, a news-oriented broadcast management system—NOBMs is implemented on top of the proposed MSE-NSS.

Jae-Gon Kim, Hyun Sung Chang, Kyeongok Kang, Munchurl Kim, Jinwoong Kim, Hyung-Myung Kim suggested a new method for summarizing a news video based on multimodal analysis of the content. The suggested method exploits the closed caption (CC) data to locate semantically meaningful highlights in a news video and speech signals in an audio stream to align the CC data with the video in a time-line. Then, the extracted highlights are described in a multilevel structure using the MPEG-7 Summarization Description Scheme (DS).

3. METHODOLOGY

3.1 Multimedia Summarization and Retrieval of News Broadcast

We define the system by identifying the 5 main components of the system; each of them will be discussed in full details in the following sections:

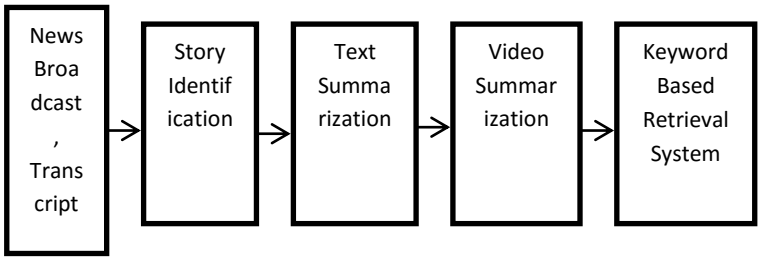

Fig-1: Main Parts of System

News Broadcast and Transcript: - Input to this system includes news Broadcast video and Transcript. News Broadcast is broadcasted on RajyaSabha TV at 9 pm as news bulletin. We downloaded news broadcast and transcript from YouTube.

Story Identification: - News Broadcast video is analyzed processed and stories boundaries are detected, news stories are identified. Anchor Person based story boundary detection is used to detect story boundaries and identify stories. The news video and transcript is segmented according to the detected news story boundary.

Text Summarization: - The most vital information condensed and extracted to produce news abstracts, i.e. a summarized passage for each news story. Natural language processing technics are used to summarize news story transcript. Lexical chain text summarization technique is used to summarize each news story.

Video Summarization: - Each news story video is summarized as per text summary generated by using lexical chain text summarization algorithm. Video processing technic are used to summarize news story. Story frames corresponding to each line in summarized story text are extracted to form news summarized story video.

Keyword based news retrieval system: - TFIDF is used to extract Important Keywords from each news stories. This could include a list of locations, persons mentioned in the news, times/dates of events, etc. Important keywords from each news story, summary of each news story are stored in a database, and a keyword based information retrieval system can be constructed. Through the use of a search engine news can be located and extracted efficiently.

3.2 System Architecture

Multimedia summary of news broadcast comprises the act of taking multimedia stream (news broadcast) comprising video, audio, and text. The news broadcast video is divided into distinct news stories segment. By using anchor person news story boundaries are identified. And start and end time of each distinct news story is recorded. By using individual story boundaries identified using anchor person corresponding news transcript is segmented into individual news stories. News transcript consisting of individual news stories is passed to TF-IDF (term frequency-inverse document frequency) and natural language processor to identify important keywords comprising names of persons, name of locations; events etc. and store it in database. Then distinct news stories are summarized using lexical chain text summarization technique. Summarized news story is stored in database. As per the textual summary generated for individual news story, individual news video/audio is also summarized. Start time of each line in news transcript is used to extract story frames in original news video. AVS file consisting summarized news story video is formed. Keyword based search is implemented to get the summary comprising text, audio/video of news story desired by user. TF-IDF weighting scheme is used as a central tool in scoring and ranking a news story's relevance given by user query.
Fig. 2 shows the overall architecture of the System.

3.3 Lexical Chain Algorithm

1. Maintains a list of interpretations.
2. Each interpretation consist of a list of lexical chains
3. Each Chain is a list of pair of nodes...
4. Each pair of node represents a link , and is in the form:
   [$word1, $line1, $word2, $line2]
5. When a link is detected, check existing chains, and
6. possible append onto chain
7. Otherwise, create new chain with the new pair.
8. Loop until reach end_line_index
9. At the end of each loop prunes the weak interpretations

Lexical Chain Text summarization algorithm is used to summarize individual news story.

There are 3 stages for constructing lexical chains:
1. Select a set of candidate words
2. For each candidate word, find an appropriate chain relying on a relatedness criterion among members of the chain
3. If it is found, insert the word in the chain and update it accordingly

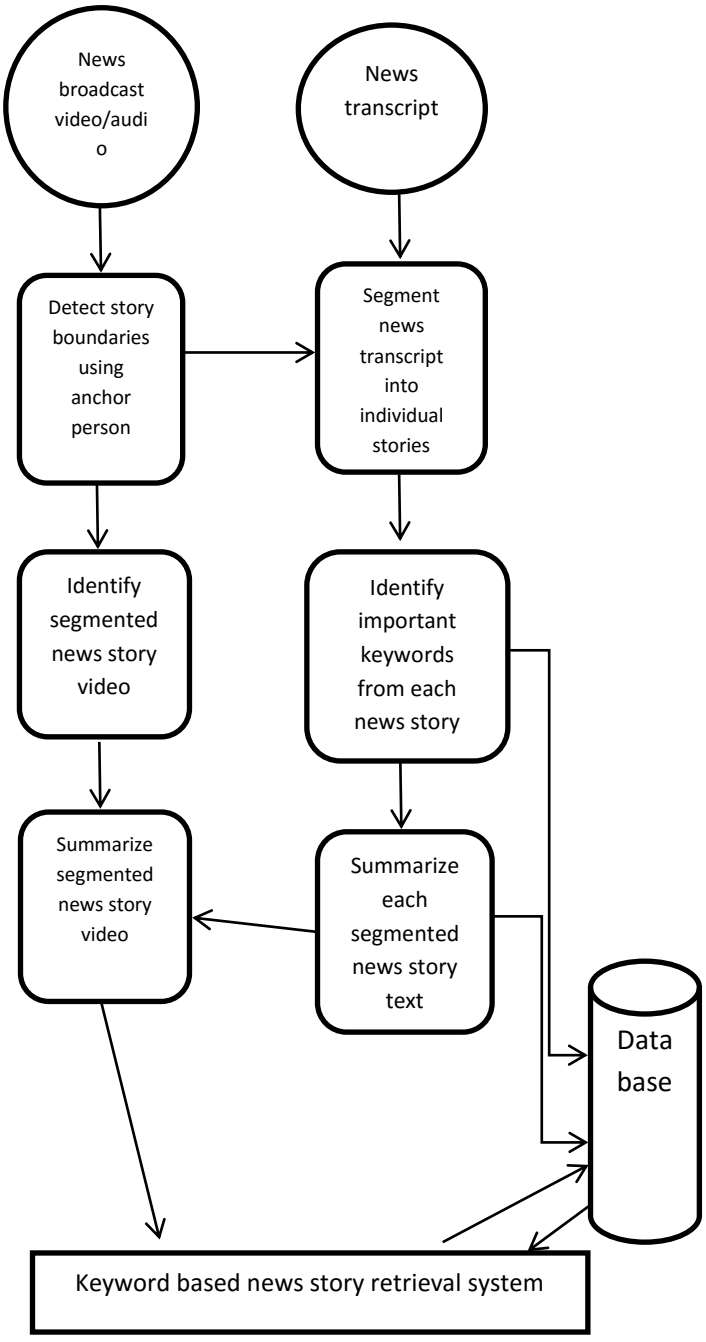

Fig.-2: Overall System Architecture

A lexical chain is created by taking a new text word and finding a related chain for it according to the "relatedness criteria". In order to be able to measure the relatedness criteria of 2 words, i.e. whether 2 words are related to each other even they are different words, a synonymy dictionary was required.

Candidate Words
All nouns in the story were chosen as candidate words for lexical chains.

Selecting Strong Chains

After all the nouns in the story have been considered, the interpretation with the highest score was chosen to represent the story. Then the 3 highest scoring chains were chosen to be the Strong chains.

Chain score=∑score generated by each link in the chain

Scores generated by each link depends on the type of link.
Extra-strong (between a word and its repetition),
Strong (between two words connected by a WordNet relation – Synonomy and Hypernomy)
Medium-strong (link between the synsets of the words is larger than one).

In selecting which chain to insert given a candidate word, extra-strong chains are preferred to strong relations, which itself is similarly preferred to medium-strong relations.

Summary
Below were the steps used for extracting summary from Strong Chains :
1. Select the representative word of the chain
2. Extract important sentences to be the summary

Representative word selected
In each chain, the word with the highest occurrence was chosen to be the representative word.

Extract important sentences.
Once the representative word was chosen, for each sentence this word appears in, we calculate the score of that function using the weighting function.

Sentence score=∑ (no. of key entity (i) detected) X weight (type of key entity (i))

4. RESULT AND ANALYSIS

4.1 Story Segmentation

Three news recordings were used for training data, and were manually processed for comparison. Each news recording was approximately 25 minutes of news broadcast. The ground truth was identified by manually viewing the video and the accompanying transcript, and these were compared to the boundaries detected by the Anchor person based story identification algorithm.

Test Results

Test 1

English news bulletin dec 09, 2017 (9 pm)

No. of stories detected-11
No. of real stories-12
False Positive-0
False Negative-1

Percentage of decision correctly made: 1 – 1 / 12 = 91.66%

Test 2

English news bulletin dec 16, 2017 (9 pm)

No. of stories detected-13
No. of real stories-15
False Positive-1
False Negative-1

Percentage of decision correctly made: 1 – 2 / 15 = 86.66%

Test 3

English news bulletin jan 06, 2018 (9 pm)

No. of stories detected-9
No. of real stories-9
False Positive-0
False Negative-0

Percentage of decision correctly made: 1 – 0 / 9= 100%

We evaluated the segmentation performance using the precision Pseg metrics.

Psg = | identified stories | - | wrong stories | / | identified stories |

Boundaries are correctly detected when a determined boundary lies within five seconds of an actual reference story boundary. Otherwise, the boundary is considered to be wrong.

Table-1: Precision

News Broadcast->	English news bulletin dec 09, 2017 (9 pm)	English news bulletin dec 16, 2017 (9 pm)	English news bulletin jan 06, 2018 (9 pm)
Precision(Psg)	0.90	0.84	1.0

The main weakness of the anchor person based story boundary detection approach seems to be the actual detection of anchor person frame. Whenever an anchor person frame has been missed, a possible story boundary will be ignored, hence bring about a drop in precision. Moreover, stories that do not start with an anchor person shot will be missed as well, which is a drawback of our anchor person based story boundary detection approach.

4.2 Story Summarization

There are no formal methods to evaluate a summarization algorithm, as 2 human generated summaries from the same passage could be very different. Therefore an intrinsic method was used to evaluate the Lexical Chain algorithm used in this project.

Story 1 – System Generated results

Beijing unveiled official emblems are for the 2022 Winter Olympics and Paralympic Winter Games the Winter Olympics emblem was inspired by the Chinese character dong which means winter the upper half of the logo was originated from the shape of a speed skater while the lower part was from the skier the emblem offer Winter Paralympics was transformed from Chinese character Fei which means flying Real Madrid will face a Brazilian club Jaime Oh in the World Cup of football a final today in last six matches rial has won four games while two matches ended in a draw on the other hand out of their six matches Gremio has a 1 - lost - and two games ended in a draw in the semi-finals Real Madrid beta al Jazeera - one while a cranial beater patro up 1-0 in the semi-final match to reach the final of the Club World Cup FIFA's ethics a watchdog provisionally

banned head of the Brazil's a Soccer Association Marco Polo del Nero for 90 days the FIFA ethics committee said that del Nero was banned from all international and domestic soccer activities and could be excluded for a further 45 days the ban was imposed after a request from FIFA's the investigatory champ chamber which is looking into unidentified violations of the organization's ethics rules that's all in this edition of fun news but before we go the Eiffel Tower is all set to attract visitors during the Christmas season with its transformation as a winter wonderland complete visitors are being welcomed by a family of a shiny penguins club chairs and replica of the Eiffel Tower with mirrors we'll leave you with these are stunning visuals thanks for watching

Story 1- Human Selected Summary

Beijing unveiled official emblems are for the 2022 Winter Olympics and Paralympic Winter Games the Winter Olympics emblem was inspired by the Chinese character dong which means winter the upper half of the logo was originated from the shape of a speed skater while the lower part was from the skier the emblem offer Winter Paralympics was transformed from Chinese character Fei which means flying Real Madrid will face a Brazilian club Jaime Oh in the World Cup of football a final today in last six matches rial has won four games while two matches ended in a draw on the other hand out of their six matches Gremio has a 1 - lost - and two games ended in a draw in the semi-finals Real Madrid beta al Jazeera - one while a cranial beater patro up 1-0 in the semi-final match to reach the final of the Club World Cup FIFA's ethics a watchdog provisionally banned head of the Brazil's a Soccer Association Marco Polo del Nero for 90 days the FIFA ethics committee said that del Nero was banned from all international and domestic soccer activities and could be excluded for a further 45 days the ban was imposed after a request from FIFA's the investigatory champ chamber which is looking into unidentified violations of the organization's ethics rules that's all in this edition of fun news but before we go the Eiffel Tower is all set to attract visitors during the Christmas season with its transformation as a winter wonderland complete visitors are being welcomed by a family of a shiny penguins club chairs and replica of the Eiffel Tower with mirrors we'll leave you with these are stunning visuals thanks for watching In this above example the extracted summary was highlighted in yellow, and the ideal summary is highlighted in green. From the above results, 8 sentences out of 15 were matched by the generated summary. However, as mentioned before, the ideal summary chosen cannot be proven to be a perfect summary, therefore the match ratio might not have any meaning to the accuracy of the system.

On the other hand, since the summary generated is an extraction summary, i.e. extracts from original content as summary, it is arguable that an accurate summary can be provided by a few sentences of the original content. The algorithm selects only 2 sentences from each strong chain. Since there were no limitations on the length of the sentences extracted, the accuracy of the results varies.

Nevertheless, extracts from the original content will have some indication to the topic of the news story, which combined together with the key entities detected, should serve as a reasonable summary for a news story.

4.3 Keyword based News Retrieval System

Detected Key Entities (important keywords) are proven to be most effective in indicating the topic of the news story. With identified Locations, organization, persons in the story, it greatly reduces the search time. E.g. (To list all the news stories of 'Narendra Modi', 'Cricket' etc.).

The search GUI was designed to be a user friendly interface, which reduces users' learning time required to adjust to the interface.

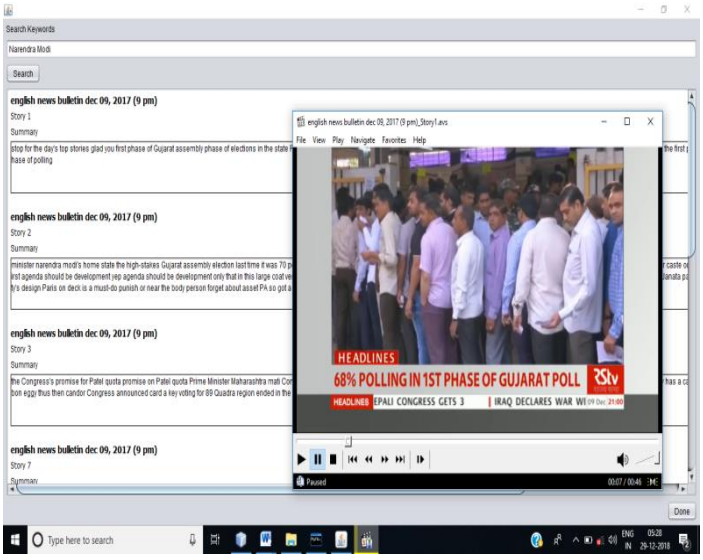

Figure-3: Keyword based News Retrieval System

5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

The most of the people give primary preference to broadcast news videos and the videos are regularly seen by millions of people in the world. Developing a system which generates extraction as well as summarization for such News video and display multimedia summary comprising video/audio and texts as per user's choice and interest is more efficient and less time consuming.

In our project, we implemented Multimedia summarization of news broadcast using anchor person based story identification and Lexical chain algorithm. Using this system one can generate multimedia summary of one or more input news broadcast and allow user to search and retrieve desired news story using keyword based search and retrieval system.

Lexical Chain summarization was implemented to provide summaries of news stories. Third party tools such as WordNet, provided text recognition abilities and vital sources of information which enabled the implementation of such algorithms.

5.2 Future Work

- We can be extended our work to identify anchor person automatically.
- This work can be extended to identify and remove advertisement from summary.
- Improvement on Summarization

The quality of the news summaries have to be improved. Investigations could be carried out on either in searching for another generic text summarization algorithm, or to improve the current algorithm. The criteria in strong chain selection should be optimized to extract more meaningful sentences.

REFERENCES

[1] Bailan Feng, Zhineng Chen, Rong Zheng, Bo Xu. Multiple style exploration for story unit segmentation of broadcast news video. Multimedia System, Springer-Verlag Berlin Heidelberg 2013

[2] Chaisorn, L., Chua, T.S., Lee, C.H.: A multi-modal approach to story segmentation for news video. World Wide Web Internet Web Inf. Systems 6(2), 187–208 (2003)

[3] Hemant Misra, Frank Hopfgartner, Anuj Goyal, P. Punitha, and Joemon M. Jose, "TV News Story Segmentation based on Semantic Coherence and Content Similarity", Dept. of Computing Science, University of Glasgow, Glasgow, G12 8QQ, UK

[4] Jae-Gon Kim, Hyun Sung Chang, Kyeongok Kang, Munchurl Kim, Jinwoong Kim, Hyung-Myung Kim. "Summarization of News Video and Its Description for Content-based Access". 2004 Wiley Periodicals, Inc

[5] Kathleen McKeown and Dragomir R. Radev, "Generating Summaries of Multiple News Articles", Department of Computer Science, Columbia University, New York, NY 10027

[6] Kuan-Yu Chen, Berlin Chen and Ea-Ee Jan, "Extractive Broadcast News Summarization Leveraging Recurrent Neural Network Language Modeling Techniques", T-ASL-04962-2014.R1.R1

[7] Mark T. Maybury and Andrew E. Merlino, "Multimedia Summaries of Broadcast News", Advanced Information Systems Center, the MITRE Corporation, 202 Burlington Road Bedford, MA 01730, USA

[8] Marcus J. Pickering, Stefan M. Rüger, "VIDEO SEARCH ENGINE USING DUAL-MEDIA SEGMENTATION", Department of Computing; Imperial College of Science, Technology and Medicine; London

[9] Mark T. Maybury, "Discourse Cues for Broadcast News Segmentation", the MITRE Corporation 202 Burlington Road Bedford, MA 01730, USA

[10] Peter Bell, Catherine Lai, Clare Llewellyn, Alexandra Birch, Mark Sinclair. "Asystem for automatic broadcast news summarisation, geolocation and translation", Centre for Speech Technology Research, University of Edinburgh, Edinburgh EH8 9AB, UK

[11] Stokes, N., Carthy, J., Smeaton, A.: SeLeCT: a lexical cohesion based news story segmentation system. J. AI Commun. 17(1), 3–12 (2004)

[12] Warren Greiff, Alex Morgan, Randall Fish, Marc Richards, Amlan Kundu, "Fine-Grained Hidden Markov Modeling for Broadcast-News Story Segmentation", MITRE

Corporation 202 Burlington Road Bedford, MA 01730-1420

[13] Wang, J.Q., Duan, L.Y., Liu, Q.S., Lu, H.Q., Jin, J.S.: A multi- modal scheme for program segmentation and representation in broadcast video streams. IEEE Trans. Multimedia 10(3), 393–408 (2008)

[14] Xie, L., Zheng, L.L., Liu, Z.H., Zhang, Y.N.: Laplacian ei- genmaps for automatic story segmentation of broadcast news. IEEE Trans. Audio Speech Language Process. 20(1), 276–289 (2012)