International Journal of Modern Trends in Engineering and Research
www.ijmter.com
e-ISSN No.:2349-9745, Date: 2-4 July, 2015

Synchronous Content Oriented Video Retrieval

Bhagwant Balasaheb Handge[1],Prof. N.R.Wankhade[2]

[1]*Department of computer Engineering, KCT's Late G.N. Sapkal College of Engineering,Anjaneri,Nashik,India, bhagwant.handge@gmail.com*

[2]*Department of computer Engineering, KCT's Late G.N. Sapkal College of Engineering,Anjaneri,Nashik,India, nileshrw_2000@yahoo.com*

Abstract-In the last decade e-lecturing has become more and more popular .Conference amount of video data in the World Wide Web (WWW) is growing rapidly. Therefore, video retrieval on the World Wide Web or a classroom a more efficient method for Video files are urgently needed. This paper presents an automated method great video indexing and searching video Conference video files. First, use the automatic segmentation and video keyframe detection, to provide a visual cues Video content for navigation. Subsequently, application metadata extract teletext OCR (OCR) for track keyframes presentation and automatic speech recognition (ASR). In the OCR and ASR Transcripts and type detection line of text on the slide using keywords extracted from these two videos and segment level Keyword extraction video browsing and content-based search. The performance and effectiveness of the proposed Indexing function is tested by evaluating Keyword extraction video browsing and content-based search.

Keywords - Lecture videos, automatic, content-based video search,video indexing, lecture video archives.

I.INTRODUCTION

It is difficult to manually index and retrieve from large video repository difficult to search with in long video clips in order to find portions of segments that the user might interested [1]. Semantic gap between low-level information extracted from the video and the user's need to meaningfully interact with it on a higher level.to find lecture in video archive rather than finding the proper position of desired keyword in video stream. The growth of e-learning video data required more efficient content based retrieval mechanism for video lectures. Is provide automatic indexing of lecture videos and find semantically appropriate part of the lecture video[2][3].

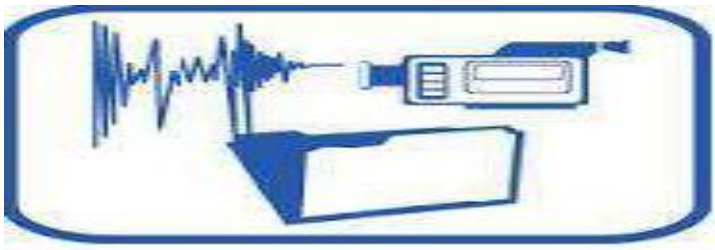


Fig 1.Content Based Video Retrieval

II. EXISTING SYSTEM

Wang et al. proposed an approach for lecture video indexing based on automated video segmentation and OCR analysis. The proposed segmentation algorithm in their work is based on the differential ratio of text and background regions. With thresholds they try to capture the slide transition. To determine the final segmentation results Synchronous shift detected key frames and associated text

books which calculate the text similarity between them as an indicator. The final segmentation result is strongly dependent on the quality of the OCR result. It might be less efficient and imply redundancies, when only poor OCR result is obtained [6].

Grcar et al. introduced VideoLectures.net in which is a digital archive for multimedia presentations.Also apply a synchronization process between the recorded lecture video and the presentation slide , which has to be provided by presenters. The slide format and the synchronization with an external document are not required [7].

Tuna et al. presented their approach for lecture video indexing and search. They segment lecture videos Key frames by using global Rahmendifferentmetrics.Then standard OCR software is used for text entry Metadata of shift streams, it to improve some image processing technology, use the OCR Score[8].

Jeong et al. proposed a lecture video segmentation method using Scale Invariant Feature Transform (SIFT) feature and the adaptive threshold in. In their work SIFT feature is used to measure the slides with similar content. Use adaptive threshold selection algorithm is used to detect transitions[9].

III. PROPOSED SYSTEM

Proposed system has an Administrator, who gives the input video to the system. After that apply ASR for Content Audio Retrieval and also identify and splitting the video into frames. Using ASR convert audio into text and also extract the segment level keywords and stored into the database. User searches the input query onto the database. If user query found then provide the result as a clustered based video.

It mainly contains following modules:

1. Capturing Frames from input video
2. Frame Classification
3. Use OCR Algorithm for Optical Character Reorganization from Each Frame of Input Video
4. Automatic Speech Reorganization for all Audio output from Input Video
5. Video and Segment Level Keywords are extracted using Output of step 3 & 4 for content based

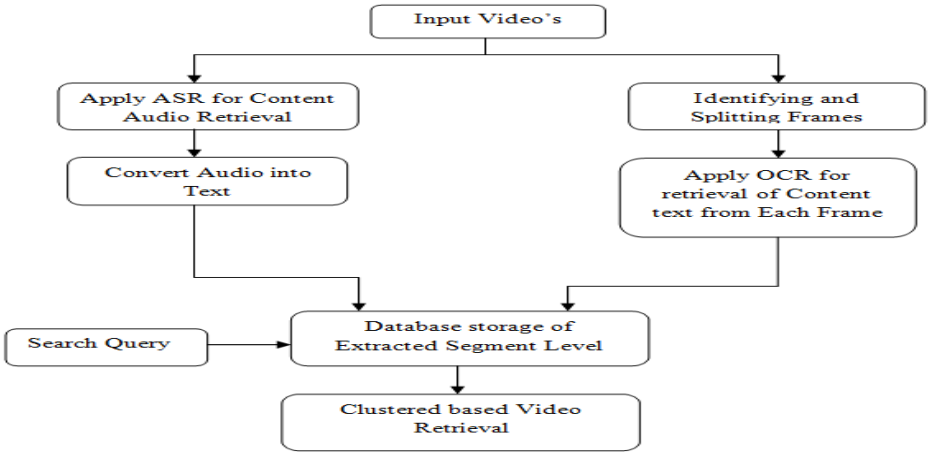


Fig 2. System Architecture

Adaptive Thresholding:

In this Heuristic algorithm method following seven steps.

1. Smooth the image by average filtering.
2. Derive the gray-level gradient magnitude.
3. Apply thresholding and a thinning algorithm to the gradient.

4. Magnitude to find the object boundary points.
5. Sample the smoothed image at the boundary points as thelocal thresholds.
6. Interpolate the threshold surface by the sampled localthresholds.
7. Segment the image by the threshold surface.
8. Remove the ghost objects by a validation process.

Video Parsing Manipulation of whole video for breakdown into key frames. Video Indexing Retrieving information about the frame for indexing in a database.Video Retrieval and browsing.

Users access the database through queries. The segmentation of the video including the border shot detection, extraction of the key image, scene segmentation and audio segmentation, extracting key frames static features, objects, audio functions and motions, exploration video data, the classification of the video and annotation, including video search interface, similarity measure, the video retrieval and relevance feedback.

Search engines are essential tools to find the desired or the most relevant information in internet. search engines can take a key advantage from using Automatic Speech Recognition (ASR) and Natural LanguageProcessing (NLP) technologies, since they allow to transcribe and enrich spoken documents, thus leading to more accurate indexes and more focused search results. To accomplish this objective, multimedia files must be processed off-line and their contents indexed to allow efficient search.

Content based video retrieval (CBVR) has wide range of applications such as consumer domain applications, quick browsing of video folder, digital museums, news event analysis, video inspection, and educational applications [4].

Key Frame Extraction: There are great redundancies among the frames in the same shot certain frames that better reflect the shot contents are selected as key frames to succinctly represent the shot. The features used for key frame extraction include colors (particularly the color histogram), edges, shapes, optical flow. Current approaches to extract key frames are classified into six categories: sequential comparison-based, global comparison-based, reference frame-based, clustering based, curve simplification based, and object/event-based.

Video Mining: A process of finding correlations and patterns previously unknown from large video databases. The task of video data mining is, using the extracted features, to find structural patterns of video contents, behavior patterns of moving entity, content characteristics of a scene, occurrence patterns and their associations, and other video semantic knowledge for intelligent video application such as reaching video retrieval.

Query and Retrieval: Once video indexes are obtained, content-based video retrieval can be performed. Upon receipt of a query is not a similarity measure method, based on the indexes to search for candidate videos according to the query. The search results can be optimized according to relevance feedback.

IV.MATHEMATICAL MODEL

Problem Description:

Let S be a technique for Retrieval of video from Database

S = { I,F,O }

Where I represents the set of inputs :

I = { D,W }

D = Set of Requirements for Retrieval of Video

W = Total Methods for retrieval of Video

F is the set of functions:

Where F = { T,F,M }

T = Data Collection from OCR

F = Data Collection from ASR
M = Threshold Comparison
O is the set of outputs:
O = { C }
C = Retrieved Video
Let,
$I_1(d)$ = Set of methods for retrieval of video.
$I_2(w)$ = Set of total methods for retrieval of video.
$I_3(t)$ = Set of data collection from OCR.
$I_4(f)$ = Set of data collection from ASR.
I(d) = F (m)U $I_1(d)$.....| ▿t ε w ε mε d |
$I_1(d)=I_2(w)U\ I_3(t)U\ I_4(f)$.....(1)
Where,
$I_1(d)=I_2(w)U\ I_3(t)U\ I_4(f)$.....from(1)
F(m) = threshold comparison.
$I_1(d)\rightarrow I_3(t)$.....(2)
$I_2(d)\rightarrow I_4(t)\rightarrow$ F(m).....(3)
$I_2(d)\rightarrow I_4(f)\rightarrow$ F(m).....(4)

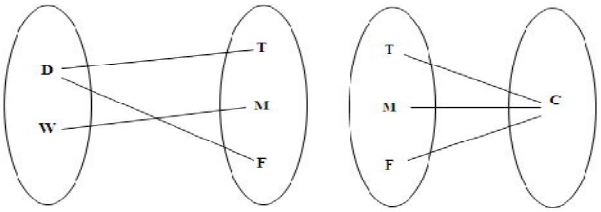
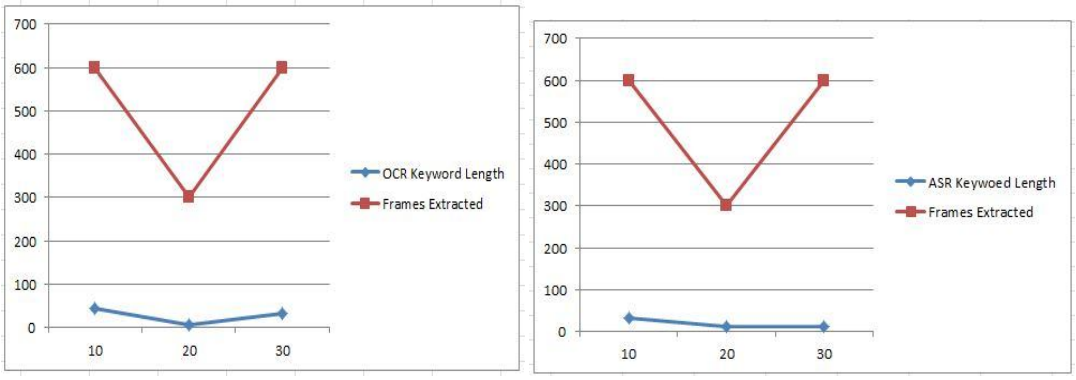


Fig 3.Venn Diagram

V. RESULT

The input video to the system. After that apply ASR for Content Audio Retrieval and also identify and splitting the video into frames. After applying the ASR convert audio into text and also extract the segment level keywords and stored into the database. User searches the input query onto the database. If user query found then provide the result as a Clustered based video.



A) OCR Keyword Length ***B) ASR Keyword Length***
Fig 4 .Segment Level Results

CONCLUSION

To manually and automatically extracted exhibit greater slide text retrieval accuracy when compared to manually and automatically transcribe spoken text. An approach to content-based lecture indexing and retrieval of video in a large lecture video archives. In order to verify the hypothesis we search the application of visual resources as well as audio from videos of the lecture Extract metadata automatically based on content. The study conducted. The usability and tool the search function will be performed in the video portal site video lecture. Automated annotation of OCR and use of open data associated resources provides an opportunity for enhance the size of the educational resources linked to a large extent. Thus more efficient search and recommendation the method can be placed in the archives lecture video.

REFERENCES

[1] Haojin Yang and ChristophMeinel, Member, IEEE, "Content Based Lecture Video Retrieval Using Speech and Video Text Information", IEEE Transactions on LearingTechnology , Vol. 7, No. 2, April-June 2014.
[2] E. Leeuwis, M. Federico, and M. Cettolo," Language modelingand transcription of the ted corpus lectures", in Proc. IEEE Int.Conf. Acoust., Speech Signal Process., pp. 232235,2003.
[3] J. Glass, T. J. Hazen, L. Hetherington, and C. Wang," Analysis and processing of lecture audio data: Preliminary investigations", in Proc. HLT-NAACL Workshop Interdisciplinary Approaches Speech Indexing Retrieval, pp. 912,2004.
[4] W. Hurst, T. Kreuzer, and M. Wiesenhutter," A qualitative studytowards using large vocabulary automatic speech recognition to index recorded presentations for search and access over the web", in Proc. IADIS Int. Conf. WWW/Internet, pp. 135143,2002.
[5] C. Munteanu, G. Penn, R. Baecker, and Y. C. Zhang,"Automatic speech recognition for webcasts: How good is good enough and what to do when it isnt", in Proc. 8th Int. Conf. Multimodal Interfaces,2006.
[6] T.-C. Pong, F. Wang, and C.-W. Ngo, "Structuring low-quality videotaped lectures for cross-reference browsing by video text analysis", J. Pattern Recog., vol. 41, no. 10, pp. 32573269, 2008.
[7] M. Grcar, D. Mladenic, and P. Kese,"Semi-automatic categorization of videos on videolectures.net", in Proc. Eur. Conf. Mach.Learn. Knowl. Discovery Databases, pp. 730733,2006.
[8] T. Tuna, J. Subhlok, L. Barker, V. Varghese, O. Johnson, and S.Shah.,"Development and evaluation of indexed captioned searchable videos for stem coursework", in Proc. 43rd ACM Tech. Symp. Comput. Sci. Educ., pp. 129-134.[Online].Available:http://doi.acm.org/10.1145/2157136.2157177.
[9] H. Sack and J. Waitelonis,"Integrating social tagging and document annotation for content-based search in multimedia data", in Proc. 1st Semantic Authoring Annotation Workshop, 2006.
[10] H. Sack and J. Waitelonis,"Integrating social tagging and document annotation for content-based search in multimedia data", in Proc. 1st Semantic Authoring Annotation Workshop, 2006.
[11] C. Meinel, F. Moritz, and M. Siebert, "Community tagging in teleteaching environments", in Proc. 2nd Int. Conf. e-Educ., e-Bus.,e- Manage. and E-Learn., 2011.
[12] S. Repp, A. Gross, and C. Meinel, "Browsing within lecture videosbased on the chain index of speech transcription", IEEE Trans.Learn.Technol., vol. 1, no. 3, pp. 145156, Jul. 2008.
[13] J. Eisenstein, R. Barzilay, and R. Davis. (2007),"Turning lecturesinto comic books using linguistically salient gestures", in Proc.22nd Nat. Conf. Artif. Intell., 1, pp. 877882. [Online]. Available:http://dl.acm.org/citation.cfm?id=1619645.1619786.
[14] J. Adcock, M. Cooper, L. Denoue, and H. Pirsiavash,"Talkminer: A lecture webcast search engine", in Proc. ACM Int. Conf. Multimedia, pp. 241250,2010.
[15] J. Nandzik, B. Litz, N. Flores Herr, A. Lohden, I. Konya, D.Baum, A. Bergholz, D. Schonfu, C. Fey, J. Osterhoff, J.Waitelonis,H. Sack, R. Kohler, and P. Ndjiki-Nya.," Contentus technologies for next generationmultimedia libraries, " MultimediaTools Appl., pp. 143, [Online].Available:http://dx.doi.org/ 10.1007/s11042-011-0971-2.

