OCRV: Optical Character Recognition from Videography

Sukanya Roy, Amit Sinha, Soumyendu Chatterjee, Avijit Das

Abstract— The goal of this paper is to provide a new methodology to detect and recognize the text from the video frames. Here the task performed is divided into three steps; combination of that detect and recognize the text from the video frame. The video frame creation devides the input video into several individual frames that grabbed and passed to the rest two phases. The text detection is a two-step approach, which involves text localization phase and the text verification phase. The text recognition involves in text verification phase and the optical character recognition phase. The final outcome of this paper is the detection of the text from the video frames in a text or word file.

Index Terms— OCR-Videography; Methodology; Video Segmentation; NTSC; Text Extraction; Text recognition; Video Frame; Binarized

---◆---

1 INTRODUCTION

Now a days digital video is very popular as an exchange medium due to large improvement in video recording and compression techniques and increasing of network-speed. Therefore audiovisual recordings are used more frequently in e-learning and e-lecturing systems. OCR from videography is a technique that can locate any text inside a digital video file via reading and automatic extraction of any notes and captions [1] that gives the actual information (like - the names of people, places or description of objects etc.) about the video being presented. Detecting the video-content requires many technologies like scanning, pre-processing, search strategies, video segmentation, feature-extraction, recognition etc. Reading the extracted notes and captions gives more appropriate information to understand the video-content. Applying OCR [3],[4],[5] on video and combining the results with various detecting techniques can improve the detection result. Although integrated character recognition in text-based videos is needed greatly.

There are various research fields which concern character recognition from videos. In character extraction from the video difficulties are present due to different conditions of title, character size and complex backgrounds. In video caption resolution of character is lower and the background complexity is more severe than in other research. The first problem is low resolution of the characters. An image size is limited by title, number of scan lines defined in the NTSC standard. Therefore, the resolution of characters in the video-caption is insufficient to implement stable and robust OCR-Videography systems. Another problem is the presence of complex backgrounds. Characters superimposed on videos often have similar hue and brightness to the background. Presence of these problems in OCR-Videography systems has opened a new area for research work. OCR [3],[4],[5] from videography is a technique that can greatly help to locate the topics

of interest in a large digital video via the automatic extraction and reading of captions and notes.

2 PROBLEM STATEMENT

Performing OCR-Videography on digital video and combining the obtained results with other video analysis techniques will improve text detection of the video content. Text recognition from digital video is a challenging problem because of the presence of rich, dynamic backgrounds, low resolution, color, etc. A strategy is required to process the video images to produce high-resolution binarized text images that resemble printed text and minimize the error rate while performing character recognition.
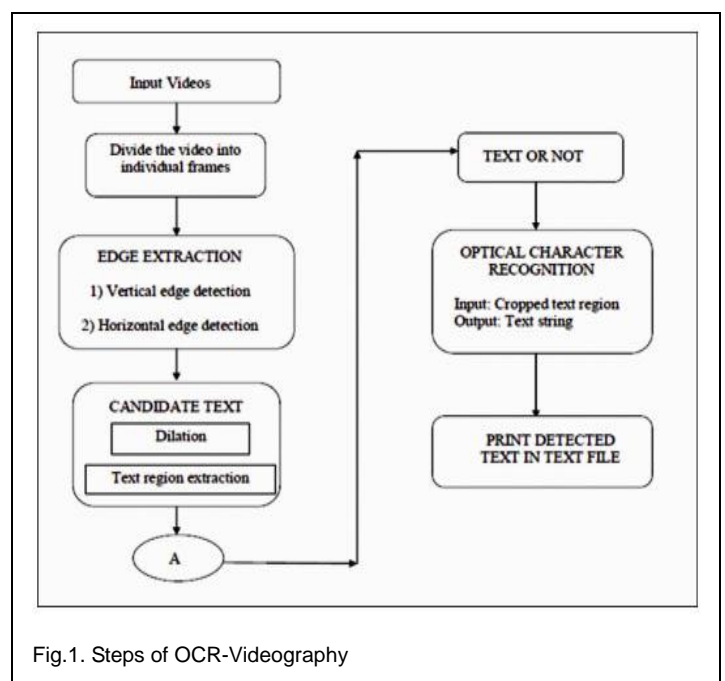
3 OCR-VIDEOGRAPHY

Fig.1. Steps of OCR-Videography

- *Sukanya Roy is a Assistant Professor of CSE dept in University of Engineering & Management, Kolkata Country-India, +91-8276973458*
- *Amit Sinha is a student of UEM, Kolkata of CSE dept 3rd year.*
- *Soumyendu Chatterjee is a student of UEM, Kolkata of CSE dept 3rd year.*
- *Avijit Das is a student of UEM, Kolkata of CSE dept 3rd year.*

The word "OCR-VIDEOGRAPHY" means performing the OCR [3],[4],[5] technique on a digital video to extract text present in the video. There are some algorithms like video-segmentation, feature-extraction, recognition that are performed to extract the text. At first the text blocks present in the digital video is detected using text detection process and then in segmentation stage these text blocks are segmented into single characters to generate the templates.

4 METHODOLOGY OF OCR-VIDEOGRAPHY

It has shown that two different types of text present in videos and images. One is scene text that appears suddenly and present in different scenes such as street-names, shop-names, T-shirts etc. Other is overlay text that is carefully directed in contrast. It is an important carrier of information and also suitable for indexing and retrieval.

Presence of different fonts with different font size, type, style, and color vary text occurrences. Some more research work is needed in these specific domains with some limited variations in these attributes.
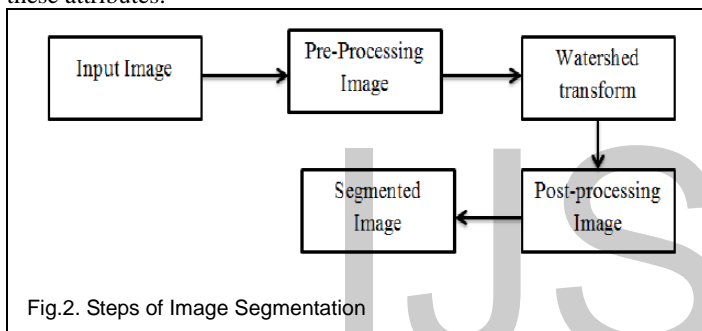

Fig.2. Steps of Image Segmentation

4.1 Image Segmentation

Image segmentation is theprocess of partitioning a digital image into multiple segments i.e. sets of pixels, also known as super-pixels. Segmentation is applyed to simplify the representation of an image into something that is more meaningful and easier to analyze. This is an initial stage of OCR-Videography.

4.2 Text Detection

There exist different types of text detection methods on the basis of color, edge and texture [2].

Color-based method: It assumes the video text is made of a uniform color. In this method the red color component is used to

obtain high contrast edges between text and background.

Edge-based method: It is used to detect overlay text that contain rich edge information. The main procedure is to apply an edge detector to the video-frames and then identify the high edge density regions.

Texture-based method: In this method the salient point detection and the wavelet transform have been used to detect the text regions.

4.3 Image Enhancement

Image Enhancement [8] is used to improve the overall quality of an image, so that the result becomes more appropiate. Sometimes same text block may present in different video-frames or background may be different for the same text block. This is a typical characteristic of text present in video-frames. The enhanced image is computed by aligning the different cases of a particular text
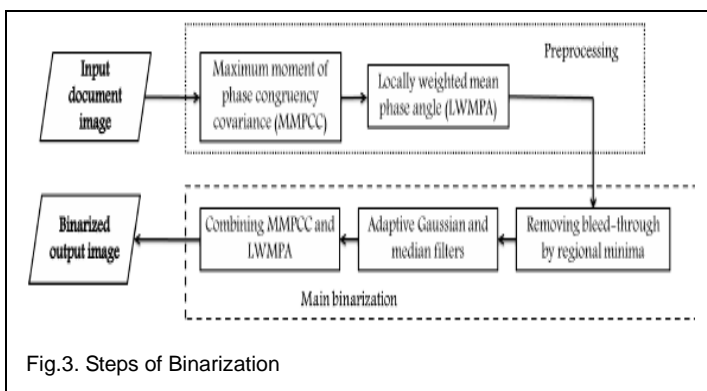

Fig.3. Steps of Binarization

region across frames and for each pixel choosing the color corresponds to the minimum intensity value across frames. Other filtration

methods like: mean, median and maximum also can be applyed.

4.4 Binarization

Pixel image can be converted into binary image using the binarization [6],[7] method. There are different techniques available for binarization; we need to choose the most appropriate one to obtain the better result.

5 APPLICATIONS OF VIDEO OCR

OCR-Videography is one of the popular research topics. We have to be more careful while creating our own OCR-Videography technique. This technique can be applied only on a snapshot of video screen or in case of a binarized image. First we have to identify the text present in video. Next we need to separate foreground and background in binarization process.

Some of the applications of OCR-Videography:

Video editing and cataloguing: catalogs video databases on basis of content relevance.

Library digitization helps to digitize cover of journals, magazines and different types of videos using advanced image and video optical character recognition (OCR).

OCR-Videography tehnique can easily mark the degraded texts present in video and increases the ability to recognize them.

Creation of "Headlines" from a live video in broadcasting media.

Object's name identification from live video. Etc.

7 CONCLUSION AND FUTURE WORK

This research is based on building an efficient system for identifying any greyscale values present in videos. The main reason of building this system is to fulfill the the needs of multimedia indexing, library digitizing and management. The technique is divided into some subtasks: recognizing the text, segment the background and binarization.

Future investigations on other aspects need to be pursued for developing video text detection and recognition applications and related multimedia retrieval and annotation applications.

One aspect is explotion of context text information of text strings from a video. Future research can focus on mining the relationship between the content of the corresponding text and categories of video shots.

Another aspect is OCR can be used for font encoding, spell checker and finding gramatical errors in sentence. A converter could be implemented to convert texts from one language to another by recognizing the characters first that are present in the video and then by converting them.

8 ACKNOWLEDGMENT

REFERENCES

[1] Sato T., T. Kanade, E. K. Hughes, and M. A. Smith, "Video OCR for
digital news archive," *in Proc. IEEE Workshop Content-Based Access*
Image Video Database, 1998, pp. 52-60.

[2] Li h., D. Doermann, 0. Kia, "Automatic Text Detection and Tracking
in Digital Video.", *IEEE Trans on Image Processing*, 2000, vol. 9, no.1, pp. 147 -156.

[3] Line Eikvil, "Optical Character Recognition", NorskRegnesentral, Oslo, Norway, Rep. 876, 1993.

[4] C.P. Sumathi, T. Santhanam, G.Gayathri Devi, "A Survey On Various Approaches Of text Extraction InImages", International Journal of Computer Science &Engineering Survey (IJCSES). Vol.3, August 2012, Page no. 27-42.

[5] Honggang Zhang, KailiZhao, Yi-ZheSong, JunGuo, "Text extraction from natural scene image: A survey",Elsevier journal on Neurocomputing ,pp.310-323, 2013.

[6] I. Pratikakis, K. Ntirogiannis, B. Gatos "Performance evaluation methodology for historical document image binarization" (IEEE).

[7] S.N. Srihari, Ying Liu "Document image binarization based on texture features" (IEEE).

[8] Gurpreet kaur, Rajdavinder Singh "Image Enhancement and Its Techniques- A Review"

[9] Mori, S,, Ching Y. S., and Kazuhiko Y., "Historical review of OCR research and development."Proceedings of the IEEE 80.7 (1992): 1029-1058.