



INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK



SPECIAL ISSUE FOR INTERNATIONAL CONFERENCE ON "INNOVATIONS IN SCIENCE & TECHNOLOGY: OPPORTUNITIES & CHALLENGES"

TEXT EXTRACTION FROM SPORTS VIDEO

PROF. VIJAY R. TRIPATHI

Department CSE, COET, Akola(MH), India

Accepted Date: 07/09/2016; Published Date: 24/09/2016

Abstract: We know that the video is consisting of sequence of images, text and audio. It is one of the sources for presenting the valuable information. Text present in video contain information which is used for automatic annotation, structuring, indexing and retrieval of video. In traditional system scores, wickets, team name, runs updated on website manually. The drawback of that system is it takes more time for updation of information on website. Another drawback of that system is some operator or receiver was needed to continuously watch the match, then they were compare the previous entered information on website with the latest information that was displayed at the bottom one-tenth part of the video. If any difference find out, then only the latest information was updated on website. In that system there is chances of errors as it is handled by human. For storing the information related to sport and updating it on the website requires database connection which introduced data management responsibility and at the same time connection problem. That's why we propose a novel method of detecting video text regions containing player information and score in sports video. First, we identify key frames from video using the Color Histogram technique to minimize the number of video frames. Then, the key images convert into gray images for the efficient text detection. We crop the text image regions in the gray image which contains the text information. Then we apply the canny edge detection algorithms for text edge detection. Using the OCR tool, the text region image is converted as ASCII text then we upload it on website using XML.

Keywords: Text extraction, Optimal Code Recognition, Canny Edge Detection, XML



PAPER-QR CODE

Corresponding Author: PROF. VIJAY R. TRIPATHI

Co Author:

Access Online On:

www.ijpret.com

How to Cite This Article:

Vijay R. Tripathi, IJPRET, 2016; Volume 5 (2): 329-336

INTRODUCTION

The proliferation of sports game broadcasting leads to the explosive growth of sports video content and the increasing need for the access of sports video content anytime, anywhere and over a variety of receiving devices. However, the rich sports video content has also resulted in much difficulty for the users to access and edit their favorite portions of sports games from huge amount of sports videos. It is clear that when accessing lengthy and voluminous sports video content, the ability to intelligently analyze sports video content to allow efficient browsing, indexing, enhancement, personalization and retrieval of sports video content is highly expected by content providers, service providers and end viewers. As there are so many video frames so it becomes problematic to decide how to index video information automatically and how to browse and manipulate them efficiently. In the past system images and video was annotated manually with a small number of key word descriptors after visual inspection by the human reviewer. It is very time consuming process. To retrieve text from video or images (Video Annotation) has become an increasingly important research area in recent years for the video information retrieval and video mining applications. Detection, localization and recognition of text captions embedded in image frames of videos are the basic steps for video retrieval and indexing. Recognizing video text information directly from video frames provides unique benefits such as (i) It is very useful for describing the contents of video sequence (ii) it can be easily extracted and compared to other semantic contents (iii) Extracted text is exactly synchronized with the image data when the event occurs (iv) It enables applications such as keyword-based image search and text-based image indexing. For example, the sports videos text information displays valuable game information such as runs, player name, team name and overs. The valuable text information extracted from sports videos is called key captions text and they can be used for video highlights. Super imposed text is mechanically added into the video frame to supplement the visual and audio content. The superimposed text is one powerful source of high-level semantics. If these text occurrences could be detected and recognized automatically, they can be a valuable source of high level semantics for indexing and retrieval. An example of the superimposed text includes headlines on TV news channels, scores display at the bottom of cricket match. The superimposed text is the most reliable clue for enable users to quickly locate their interested content in an enormous quantity of video data, many efforts have been taken for video indexing and summarization. There are basically three reasons: 1) it is closely related to the current content of video 2) It has distinctive visual characteristic and 3) the state of art optical character recognition (OCR) techniques are far more robust than the existing speech analysis techniques and visual object analysis.

II. MAIN CONCEPTS

Text extraction in video consists in three steps. The first one is to find text region in original images. Then the text needs to be separated from background. And finally a binary image has to be produced (for example, text is white and background is black) Difficulties of such a project can be classified in following main categories: 1. Background and text may be ambiguous. 2. Text color may change: text can have arbitrary and non-uniform color. 3. Background and text are sometimes reversed. 4. Text may move. 5. Unknown text size, position, orientation, and layout: captions lack the structure usually associated documents.6. Unconstrained background: the background can have colors similar to the text color. The background may include streaks that appear very similar to character strokes. 7. Color bleeding: lossy video compression may cause colors to run together. 8. Low contrast: low bit-rate video compression can cause loss of contrast.

III. BACKGROUND

From observation on TV program, normally the superimposed text is displayed only a single line which appears in one-tenth part of image. The super imposed text generally divided into two types that are (i) Moving text in the video frame. i.e., the advertisement messages/text and flash news may be moving from left to right or right to left. (ii)Nonmoving Text in a video frame, i.e, the sports score card display without moving.The text detection algorithm can be classified in two categories (i) those working on the compressed domain(ii)those working on the spatial domain. Compressed domainincludes both in the compressed and in the semi-compressed domain.It is based on the localization of static characters over moving background where the text does not occupy the whole frame and that it has to appear at least in three frames. Spatial domainmethods work with the pixel values and positions.The text detection problem involves locating regions in a video frame that contain text. It refers to the determinationof the presence of text in a given frame normally text detection is used for a sequence of images.Text Region Extraction is the process of determining the location of text in the image. The video text detection is based on the special characteristics of the text such as contrast, color, font size, font shape etc.

a) Size:-

The text size considers depending on the application domain.

b) Alignment:- The characters in the superimposed text lie horizontally.

c) Inter-character distance:-

The characters in a text line have a uniform distance between them.

d) Color:-

The characters in a text line tend to have similar colors. However, video images contain text strings with more than two colors for effective visualization.

e) Edge:-

Most super imposed and scene texts are designed to be easily read, thereby resulting strong edges at the boundaries of text and background.

IV. METHODOLOGY

In this paper, we provide a robust detection method for the superimposed text in sports videos. Following are the six steps that we used for extraction of superimposed text in a video frame: (i) Video Frame Extraction (ii) Key Frame Extraction (iii) Grayscale Conversion (iv) Cropping the Video Image (v) Canny Edge Detection (vi) Text region Retrieval. The video is made up of sequence of images that are nothing but video frames. In the first step, we extract all frames in the video and saved as JPEG images. In the second step, we reduce the number of frames by taken key frame only for that we used the Histogram technique. Key frame is the frame which has major difference from previous frame. It uses adjacent frame subtraction with threshold value. The next process is grayscale conversion. We convert the color image into gray scale image by using binarization method. The images which are converted into gray scale are easy to process. Then next step is to crop the images into one tenth of its original size. Because of the score cards in the sports video are visible only in bottom of frame. The region consists of score card which contains runs, wickets, player name, overs, team name etc. The other regions of the image are not useful. So, we discard remaining part of the images. The processing time and memory requirement reduced in cropped image process. Then we apply the Canny Edge Detection implement and more efficient as compared to other algorithms. From the edge detected images, text region can be identified. The flowchart of the proposed method is illustrated in Fig. 1. The extracted regions can be used as input to an Optimal character Recognition (OCR) algorithm for the extraction of the text region in the cropped image. The canny edge detection algorithm is easy to implement and more efficient as compared to other algorithms. From the edge detected images, text region can be identified. The flowchart of the proposed method is illustrated in Fig. 1. The extracted regions can be used as input to an Optimal character Recognition (OCR).

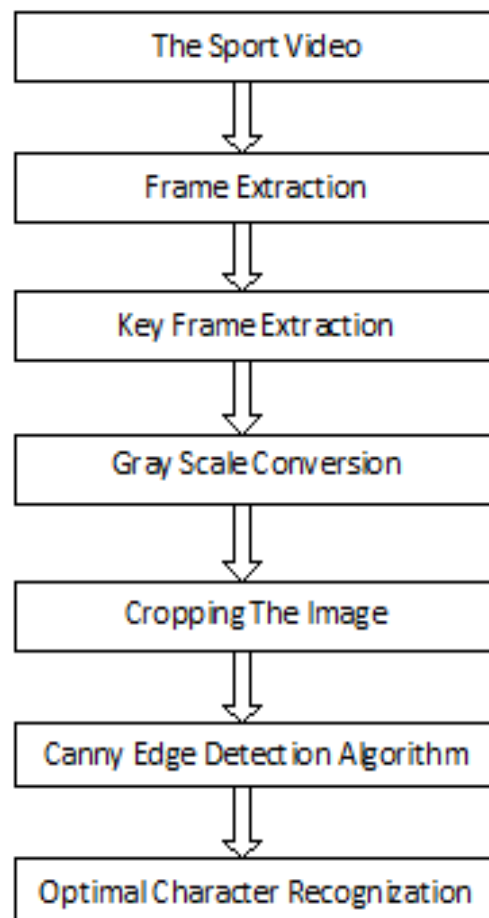


Fig.1 Steps of text extraction from sport video

By using optimal character recognition technique we get our desired output that is score card and that score card we upload on website by using XML.



a) Original Image Taken From Video

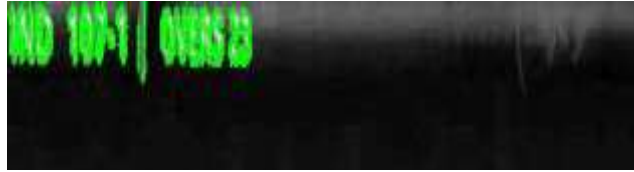
Then key frame images are converted as Grayscale images for easy process.



b) Gray Scale Image



c) Cropped Image



d) Canny Edge Detection

Finally, the super imposed text region extracted and is shown in Fig



Fig : Text Region

The extracted text images can be transformed into ASCII text using OCR tool. The ASCII text is used to video indexing and retrieve the video.

V. CONCLUSION

There are numerous application of text extraction from images like liscence plate extraction, video content analysis. We made one application which can extract the desired text from sport video like runs, overs, player name, team name, wickets etc and display it on website within a short period of time using XML. This application can be called as robust method because there is no chances for human error as it is automatically update the score card. In future research, we have to extract only the play shots. We can the audio information such as applause and cheering

To generate more exciting sports summaries.

REFERENCES

1. Keechul Jung, KwangIn Kim, and Anil K. Jain. "Text information extraction in images and video: a survey," International Journal of Pattern Recognition, Volume: 37,Number:5, pp:977 997, 2004.
2. Luo B., Tang X., Liu J., and Zhang H.-J., "Video caption detection and extraction using temporal information," In ICIP'03, pages 297-300, 2003, Barcelona, Spain, September.

3. David Crandall, Sameer Antani, Rangachar Kasturi, "Extraction of special effects caption text events from digital video," International journal on document analysis and recognition, Volume: 5, pp:138- 157, 2003.
4. Miriam León, Antoni Gasull, "Text Detection In Images And Video Sequences," Source: http://gpstsc.upc.es/imatge/pub/ps/IASTED05_Leon_allo_Gasull.pdf
5. Christian Wolf and Jean-Michel Jolion, "Extraction and recognition of artificial text in multimedia documents," Pattern Analysis and Applications (PAA), Volume 6, Number 4, February 2004 , pp. 309-326(18), Springer
6. Malobabić, Jovanka and O'Connor, Noel E. and Murphy, Noel and Marlow, Seán "Automatic detection and extraction of artificial text in video," In: WIAMIS 2004 - 5th International Workshop on Image Analysis for Multimedia Interactive Services, pp:21-23 April 2004, Lisbon, Portugal.