Text extraction from cricket video- Comparative study of algorithms

Ms. Divya L. Madan; Ms. Maithili. S. Puranik ; Ms. Mohini. N. Nahate ; Ms. Ragini. N. Nahate & Asst. Prof. Kapil. O. Gupta

ABSTRACT

Video is one of the sources for presenting the valuable information. It contains sequence of video images & text information. Text data present in video contains useful information for automatic annotation, structuring, mining, indexing & retrieval of video. Nowadays manually added text in video sequences provides useful information about their contents. It provides supplemental but important information for video indexing & retrieval. To update the information on the website every one operator is required for watching and feeding data. To overcome these flaws we are using certain algorithms namely, Sobel algorithm, Prewitt's algorithm, Robert's algorithm & Canny edge detection algorithm. We are studying the comparison between these algorithms. This project contains firstly the processing of the video then we identified key frames from minimize the number of video frames. Then the key images are converted into gray images for the efficient text detection. Generally, the superimposed text displayed in bottom part of the images in the cricket video. So, we cropped the text image regions in the gray image which contains the text information. Then we are applying the above mentioned algorithms for text extraction from the cricket video.The algorithm that takes less time is the most efficient to work with. Hence, there is no need of any operator to feed the data on the website.

Keywords: Image Processing; text extraction; text recognition; text in video; video retrieval; video annotation; cricket video; video summarization; video text information and superimposed text.

Introduction

Increasing availability of video data has rekindled interest in the problems of how to index video information automatically and how to browse and manipulate them efficiently. Traditionally, the images and video sequences have been manually annotated with a small number of keywords descriptors after visual inspection by the human reviewer. This process can be very time consuming. Text information retrieval from video images (video annotation) has become an increasingly important research area in recent years for the video information retrieval and video mining applications. Detection and recognition of text captions embedded in image frames of videos is an important component for video retrieval and indexing. Video text detection and extraction is an important step for information retrieval and indexing in video images. Recognizing video text information directly from video provides unique benefits[1]. They are:

1. It is very useful for describing the contents of video sequence.
2. It can be easily extracted and compared to other semantic contents.
3. Extracted text is exactly synchronized with the image data when the event occurs.
4. Manual login may not be feasible for large collection of archived videos.
5. It enables applications such as keyword-based image search,

automatic video logging, and text-based image indexing.

6. It is an important component for the automatic annotation, indexing and parsing of images and videos.

For example:

The cricket videos text information displays valuable game information, such as scores and players. The valuable text information extracted from cricket video is called a key caption text, and they can be used for video highlights or content search.

The objectives of the project are:

- ○ We compare the Sobel's algorithm, Prewitt's algorithm, Robert's algorithms and canny edge detection algorithm to find out the best one for text extraction from the cricket video.

- ○ The algorithms that takes less time for extracting the text from the video is the most efficient to work with.

- ○ Feeding data automatically to the website is the major objective of the project.

Literature Survey

Luo B. Tang X., Liu J., and Zhang H.-J[1]- Studied the detail knowledge about the canny edge detection algorithm. Also the procedure of processing the video and framing it. We have also studied the details of gray scale algorithm & extracting the text from the video.

Cheolkon Jung and Joongkyu Kim[2]- Studied the details about the various algorithms that is the Sobel's algorithm, Prewitt's algorithm, Robert's algorithm and Canny edge detection algorithm, also the details about the comparisons between them and the variations.

David Crandall, Sameer Antani, Rangachar Kasturi[3]- Here comparison and analysis of several kinds of classical algorithms of image edge detection, including Robert's, Sobel's, Prewitt's and Canny edge detection algorithms.

Architecture

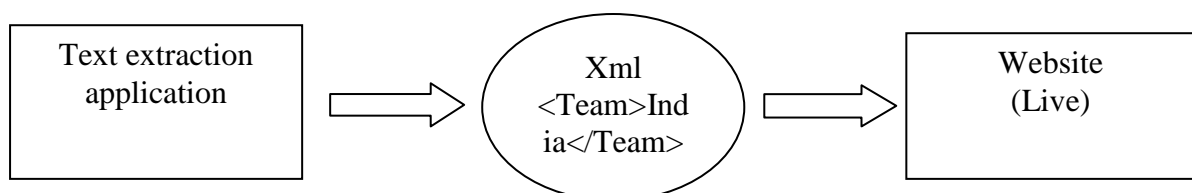Text extraction application → Xml <Team>India</Team> → Website (Live)

Fig 3.1 database architecture

Here, two machines are used for the process of extraction. That is the text will be extracted from one machine. After extraction the extracted data will be stored into the xml schema from where the xml parser will parse the data and will be displayed on the other machine where the live website is launched. And an xml parser is nothing but a system to read or scan the xml files. Here no database is used for the storage of data instead we are using an xml schema.

Here, a robust detection method for superimposed text in cricket videos is provided. An efficient algorithm for super imposed text region extraction of super imposed text in video is presented. Comparison of certain algorithms is done here to study the most efficient algorithm from them. The algorithms that are being used are: Sobel's algorithm, Prewitt's algorithm, Robert's algorithm and Canny edge detection algorithm respectively. The algorithm that gives the most efficient and fast results is the one that is best worked with. Here, framing, cropping and gray scaling methods are also performed to extract the text from the cricket video. All these methods are performed first so that after extracting text the algorithms can be applied.

The application is run on the MATLAB software. Using a file called "mainClass.m" we can run as well as see the coding of the page. There is one more file called "mainClass.fig" where you can design the desired window you wish to display. So far first module is designed where one text area and two pushbuttons can be seen in the window using which you can browse the video and process it.

The comparisons between the algorithms are as follows: Sobel operator is in the form of the filtering operator. It is used to extract the edge. Each point on the image is the 2 nuclear convolutions.

One checks maximum response of the vertical edge & the other one checks maximum response of the horizontal edge. Robert's operator is the first operator which uses a partial differential operator to find the edge. It uses the approximation between the 2 adjacent pixels of the diagonal direction of the gradient amplitude to detect the edge. The Prewitt operator is one type of an edge model operator. Model operator is made from the ideal edge sub-image composition. Detect the image using the edge model one by one, and take the maximum value of the model operator that is most similar to the detected region as the output of the operator.
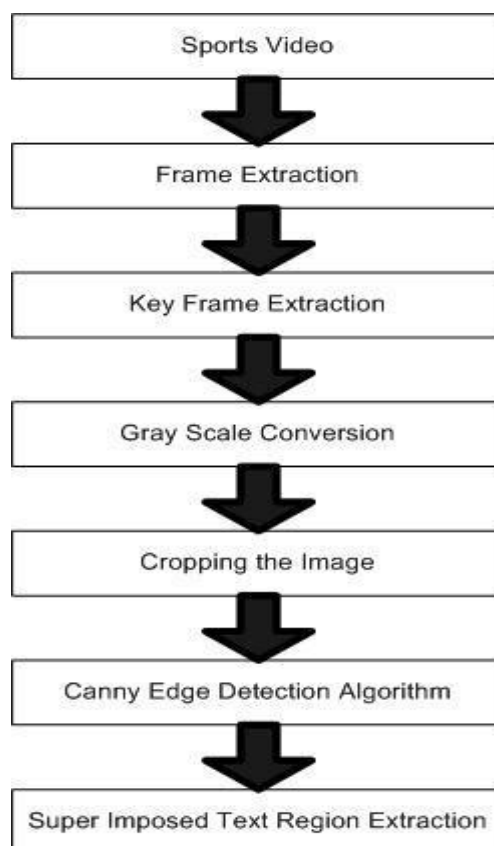

Fig 3.2 flow chart of procedure

SOBEL ALGORITHM

- Sobel operator is in the form of the filtering operator.
- It is used to extract the edge.
- Each point on the image is the 2 nuclear convolutions.
- One checks maximum response of the vertical edge & the other one checks maximum response of the horizontal edge.
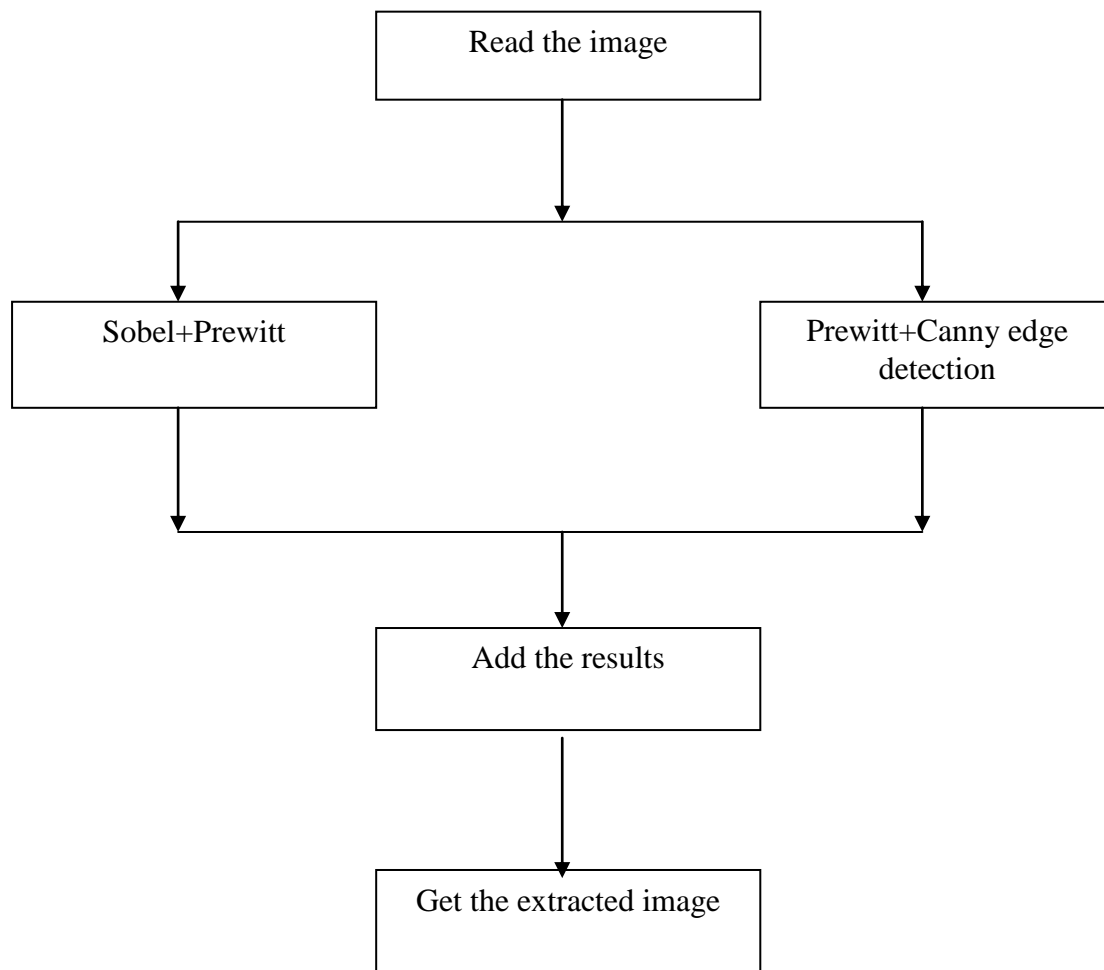
```
          ┌─────────────────────┐
          │   Read the image    │
          └─────────────────────┘
                    │
        ┌───────────┴───────────┐
        ▼                       ▼
┌─────────────────┐   ┌─────────────────────┐
│  Sobel+Prewitt  │   │ Prewitt+Canny edge  │
│                 │   │     detection       │
└─────────────────┘   └─────────────────────┘
        │                       │
        └───────────┬───────────┘
                    ▼
          ┌─────────────────────┐
          │   Add the results   │
          └─────────────────────┘
                    │
                    ▼
          ┌─────────────────────┐
          │ Get the extracted   │
          │       image         │
          └─────────────────────┘
```

Fig 3.3 flowchart of sobel, prewitt & canny edge detection algorithm

ROBERT'S OPERATOR

- Robert's operator is the first operator which uses a partial differential operator to find the edge.
- It uses the approximation between the 2 adjacent pixels of the diagonal direction of the gradient amplitude to detect the edge.
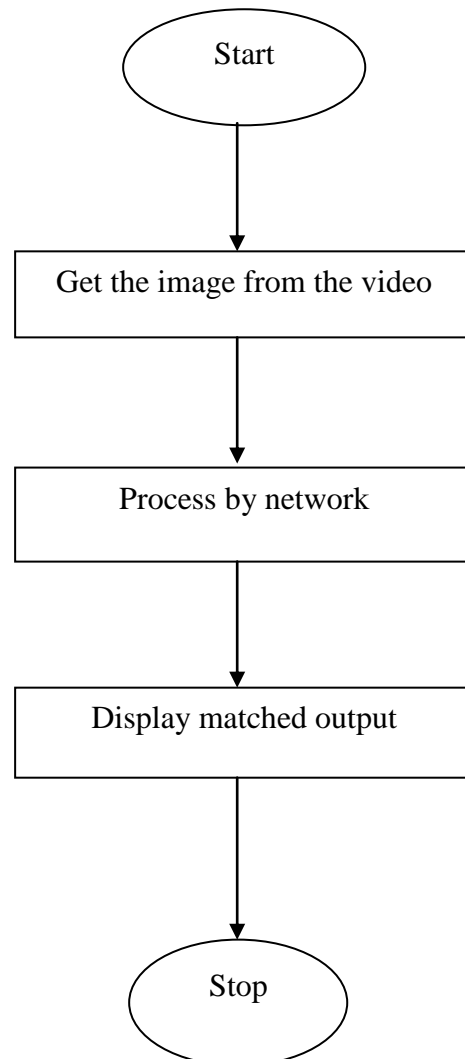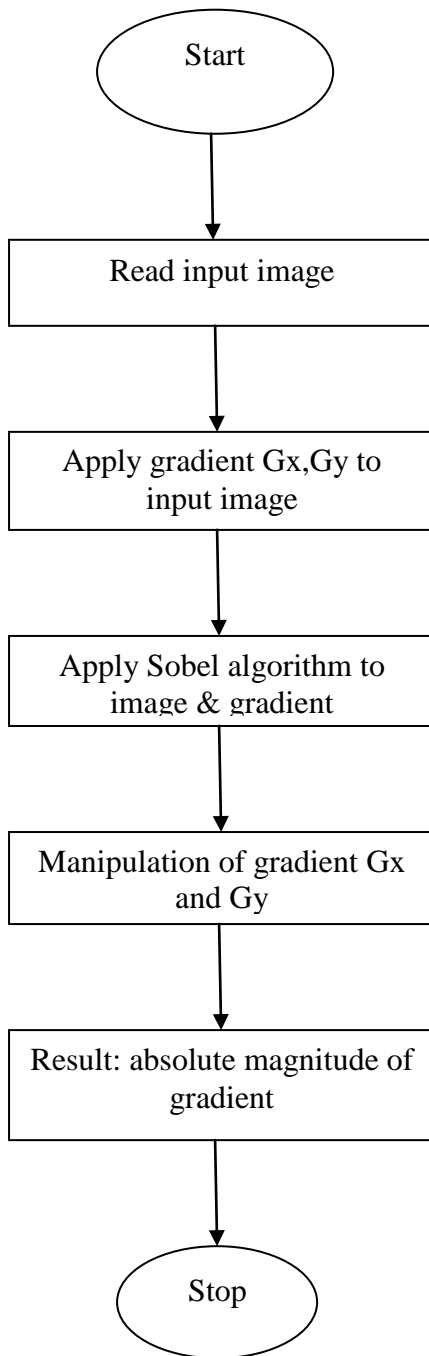
```
                    ┌─────────────┐
                   (    Start      )
                    └──────┬──────┘
                           │
                           ▼
             ┌─────────────────────────────┐
             │  Get the image from the video │
             └──────────────┬──────────────┘
                            │
                            ▼
             ┌─────────────────────────────┐
             │       Process by network     │
             └──────────────┬──────────────┘
                            │
                            ▼
             ┌─────────────────────────────┐
             │     Display matched output   │
             └──────────────┬──────────────┘
                            │
                            ▼
                    ┌─────────────┐
                   (     Stop      )
                    └─────────────┘
```

Fig. 3.4 Robert's flowchart

PREWITT OPERATOR

- The Prewitt operator is one type of an edge model operator.
- Model operator is made from the ideal edge sub-image composition.

```
        ( Start )
            │
            ▼
 ┌──────────────────────┐
 │   Read input image   │
 └──────────────────────┘
            │
            ▼
 ┌──────────────────────┐
 │ Apply gradient Gx,Gy to │
 │     input image      │
 └──────────────────────┘
            │
            ▼
 ┌──────────────────────┐
 │ Apply Sobel algorithm to │
 │   image & gradient   │
 └──────────────────────┘
            │
            ▼
 ┌──────────────────────┐
 │ Manipulation of gradient Gx │
 │       and Gy         │
 └──────────────────────┘
            │
            ▼
 ┌──────────────────────┐
 │ Result: absolute magnitude of │
 │      gradient        │
 └──────────────────────┘
            │
            ▼
        ( Stop )
```

○ Detect the image using the edge model one by one, and take the maximum value of the model operator that is most similar to the detected region as the output of the operator.

CANNY OPERATOR

- The canny edge algorithm runs in 5 separate steps:
1. Smoothing: - blurring of the image to remove noise.
2. Finding gradient: - the edge should be marked where the gradients of the image large magnitude.
3. Non maximum suppression: - only local maxima should be marked as edges.
4. Double thresholding: - potential edges are determined by the thresholding.
5. Edge tracking by hysteresis: - final edges are determined by suppressing all the edges that are not connected to a very certain (strong) edges.

Input image

Pre-processing

Media filter

Sobel operation

Post-processi

Thinning
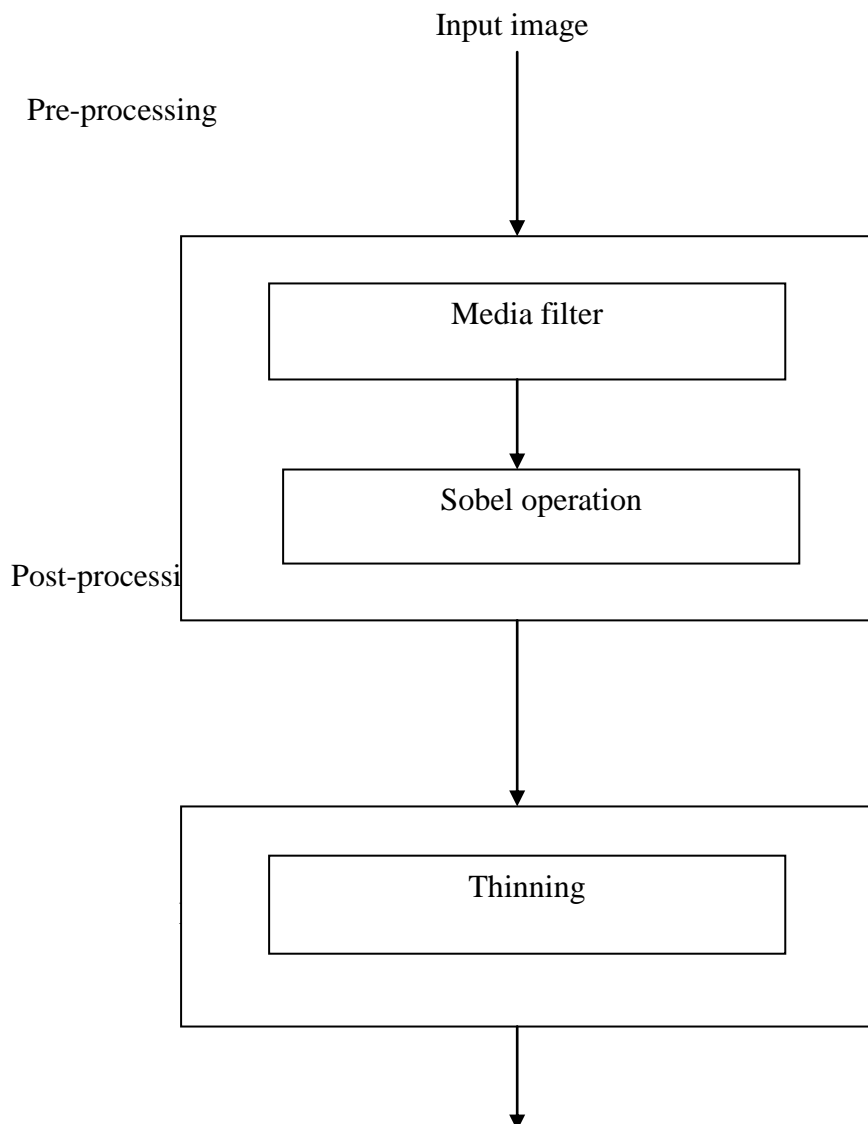
COMPARISON OF EDGE DETECTION

to evaluate these algorithms by setting different thresholds.

The result of Roberts operator:
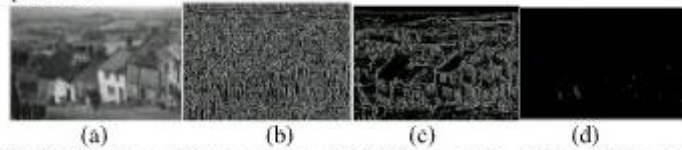
..(a) original image (b) threshold = 0.00 (c) threshold = 0.05 (d) threshold = 0.3

The result of Sobel operator:
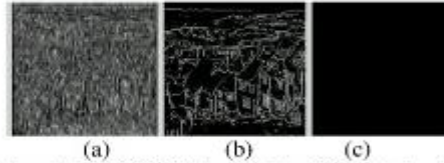
(a) threshold =0.00 (b) threshold = 0.05 (c) threshold = 0.3

The result of Prewitt operator:
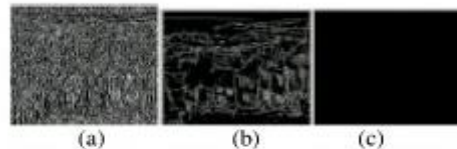
(a) threshold = 0.00 (b) threshold = 0.05 (c) threshold = 0.3

The result of Canny operator:
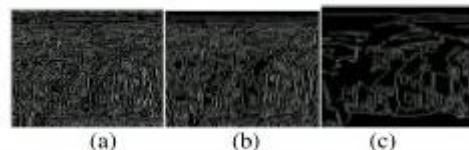
(a) threshold = 0.00 (b) threshold = 0.05 (c) threshold =0.3

Fig 3.7 comparison of algorithms

Working Details

Tools of Matlab

GUI (also known as a graphical user interfaces) provide point and click control of software application need to learn a language or a type commands in order to run the application.
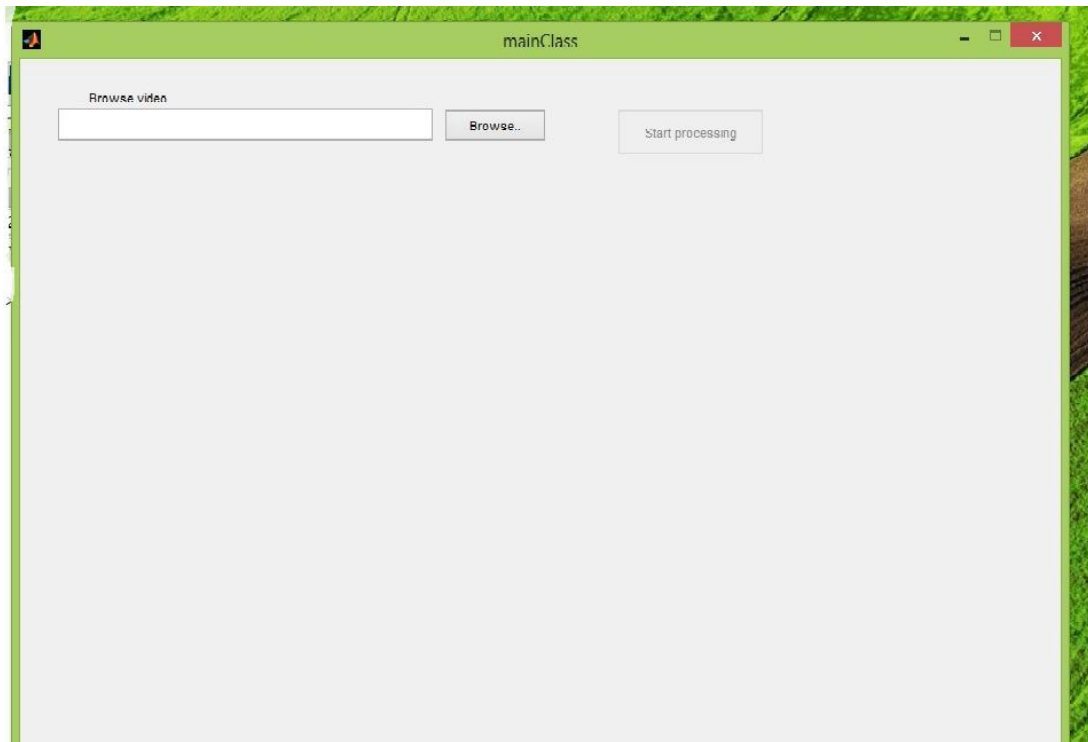
Fig 4.1 screenshot of mainClass of browse window

Step 1:
Open MATLAB.
Run the file named "mainClass.m", a window appears to browse the cricket video.
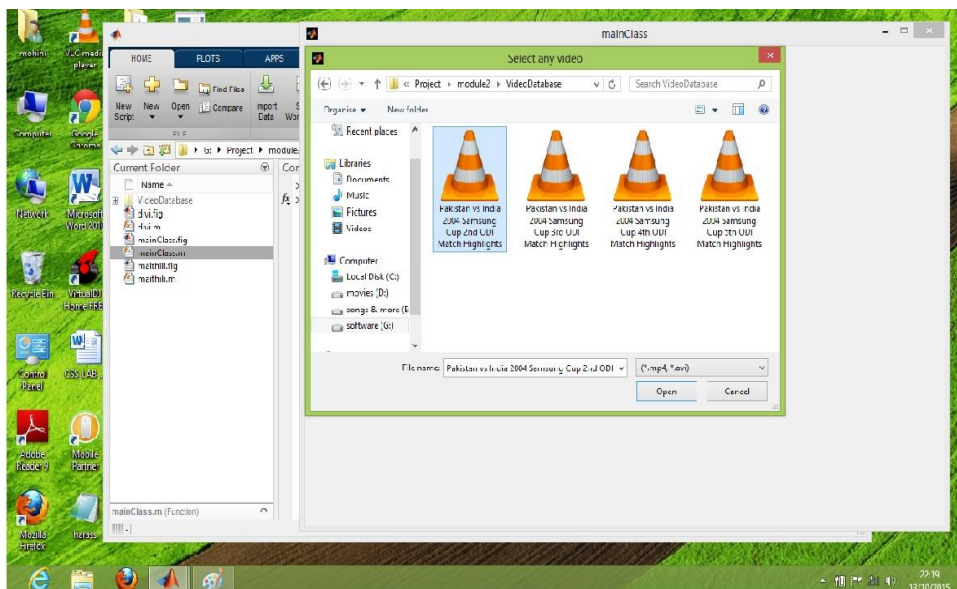

Fig 4.2 screenshot of selecting any video

Step 2:

Click on the "Browse" pushbutton to select the cricket video.

Select the video from the respective folder and click on open button.

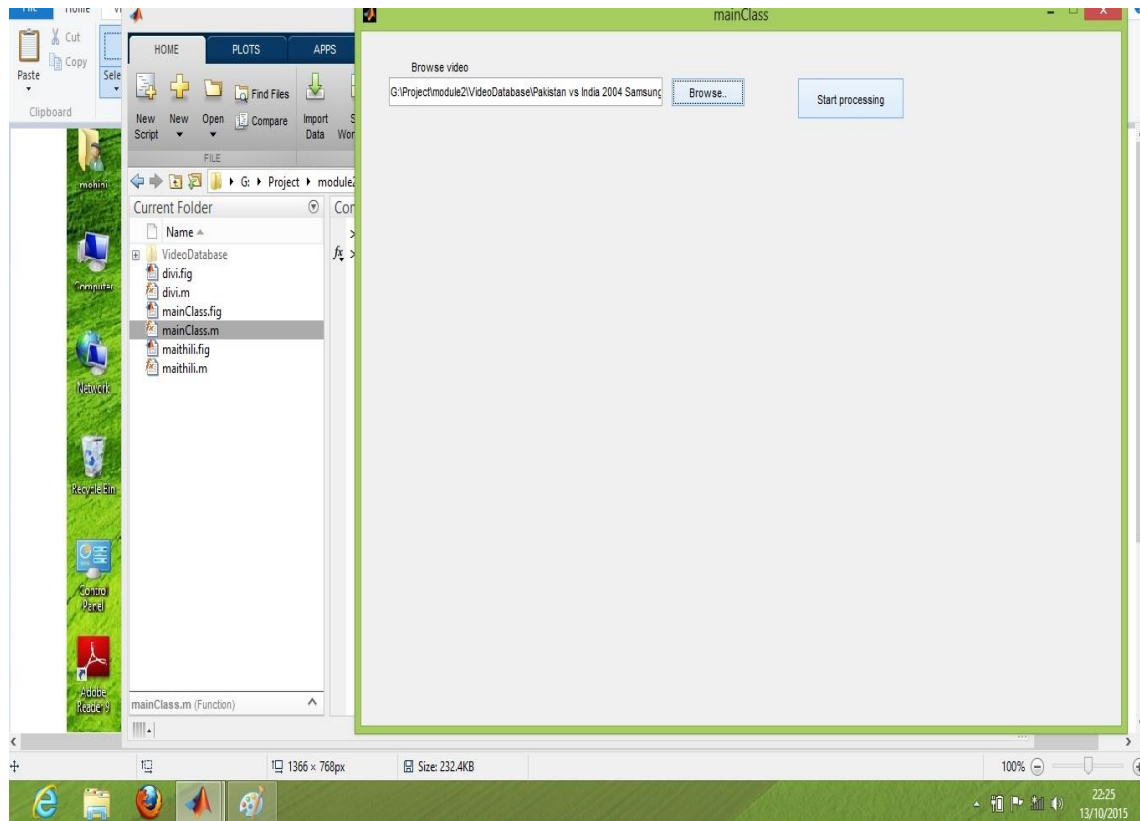The video must be with the extension of either ".mp4" or ".avi".

Fig 4.3 screenshot of browse the selected video

Step 3:

After selecting the video from the folder and opening it, the url of the video appears in the text area. Click on the other pushbutton "Start Processing" to process the cricket video.
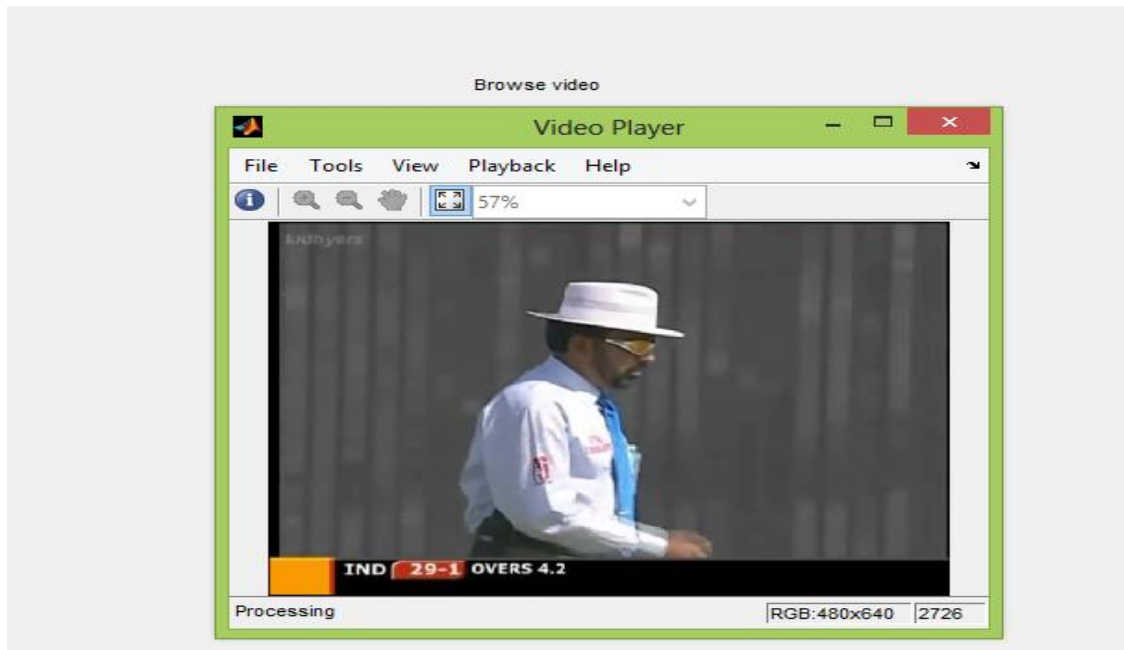
Fig 4.4 screenshot of playing video

Step 4:

After processing the video it runs successfully.

To have better view of the video click the "fit to window" from the tools menu bar..

Advantages And Disadvantages

5.1 Advantages:

Following are the advantages:-

- O The updating speed of the data on website is very fast.

- O There is no scope of human errors in system.

- O No need to use operator.

- O It is very robust.

- O Data is automatically updated.

- O Less time is required for updating data.

- O No database is required for storage of information.

5.2 Disadvantages:

- Without internet connection there is no access to the website.

- The application will not work without the xml parser.

- Xml parser is used as we are using xml schema only.

- Video should be of extension .mp4 or .avi.

Application And Future Scenario

There are numerous applications in the following, we briefly describe some of these applications.

- Wearable or portable computers: with the rapid development of computer hardware technology, wearable computers are now a reality.

- Content-based video coding: The. mp4, .avi standard supports video.

- Text-based image indexing: This involves automatic text-based video extracting methods using key caption.

The future scenario of this project is to implement framing on the processed video and apply the algorithms. We need to compare the algorithms and find out the best working one so that the errors are reduced and the website works in robust environment. Also we have to study the comparison in the upcoming modules.
The xml parser will parse the data and will be displayed on the other machine where the live website is launched. And an xml parser is nothing but a system to read or scan the xml files. Here no database is used for the storage of data instead we are using an xml schema

Conclusion

Our system makes automatic changes in the website with the robust methods. Comparison of algorithms is studied practically. We can reduce time for retrieval of database using xml parser for displaying information on website with automatic website refreshing. We made one application which can extract the desired text from the cricket video like runs, over, 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 a robust method because there is no chances for human errors as it is automatically updated the score card. Our proposed algorithm can only detect the videotext in the boundary of the image. In general, there is a little valuable information in non-play shots because these non-play shots include scene then the game is temporarily halted by scoring, no run play, or time-out, and do not carry any meanings related to the contents of a game. The text

extraction results are inappropriate for general OCR software text enhancement is needed for low quality video images and more adaptability is required for general cases.

Video text has some temporal features:

> 1) The same text usually appears over several continuous frames.

> 2) Motive characters have a dominant translation direction: linearly horizontal or vertical direction.

> 3) When text appears or disappears, text region colour change significantly in adjacent frames.

Reference:

[1] Luo B. Tang X., Liu J., and Zhang H.-J(2003) *"Video caption detection and extraction using temporal information,"* In ICIP'03, pages 297-300, Barcelona, Spain.

[2] Cheolkon Jung and Joongkyu Kim(2008) *"A Novel Approach for Key Caption Detection in Golf Videos Using Color Patterns,"* ETRI Journal, Volume 30, Number 5, pp-750- 753.

[3] David Crandall, Sameer Antani, Rangachar Kasturi(2003) *"Extraction of special effects caption text events from digital video,"* International journal on document analysis and recognition, Volume: 5, pp:138- 157.