International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P), Volume-3, Issue-7, July 2015

Semi-Automated Car Surveillance System using Information Retrieval

Tejashree Gharat, Tejasvi Dhodi

***Abstract*— With the advent of surveillance video systems, security professionals now face a challenge to determine the interesting portion of the video. The core benefit of converting a relentless deluge of data into actionable information that can shape strategies and improve processes is to maximize the value of raw uncut surveillance video footage data through rigorous analysis that reveals key information related to security. Accidents are increasing, vehicle thefts occur often, other than these people hide stolen items in their cars and might change their number plates. Keeping the various car related security concepts in mind, we are proposing a system that deals with parking lot entry and exit video footage. Our system processes this footage to achieve security and managerial goals. The content upon which our system relies is number plate and structure of the car that are extracted by using techniques in Image and Video Processing like Morphological Processing, Training Algorithms and Speed up Robust Features Detection.**

***Index Terms*— SURF, CBVR, CBIR, XML, Matlab GUIDE**

I. INTRODUCTION

Often it becomes hard to find the appropriate video content you are trying to search over the web; or retrieving a particular portion of the video which is of interest. **Content Based Video Retrieval (CBVR)** is a way to simplify and speed up accurate access to video data. The advances in technology such as capturing, refining and transferring video content has advanced over the years, but still there is a lack of efficiency for retrieving content based video data. It requires more than just connecting to video databases and fetching the information to the users through networks.

A wide range of CCTV security cameras that enable you to protect, secure and ensure the safety of your housing or residential society or townships or remotely located unmanned sites or facilities are available in market. The CCTV Security Cameras can be installed in parking lots to monitor your cars from theft, in corridors and lobby for monitoring unauthorized visitors, in elevators to prevent vandalism and also in the play area of your complex to monitor your children. People hardly have enough time to go through the entire video footage. Users are only interested in the portion where some activity takes place. So, to reduce users' hardship we propose a semi-automated system that provides us with various security concepts.

The proposed system will use CBVR, CBIR, Image Data Mining and Processing algorithms to determine the interesting portion of the video based on the user's query. This paper shows how different algorithms can be integrated to create a system that will overcome the commonly faced problems in surveillance.

Tejashree Gharat, Bachelor of Engineering in Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, University of Mumbai.

Tejasvi Dhodi, Bachelor of Engineering in Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, University of Mumbai

II. THE PROPOSED SYSTEM

A. Objectives

- Number plates of all the cars entering and exiting the parking lot are recorded in a database by extracting them from the video frames.
- The Structure of the car is detected and stored in the database.
- The frame that best describes the car from its back and side view are retrieved to recognize the car.
- A section of the video from the entire video can be extracted based on the user's requirement like the day, the gate (entry or exit).
- If the car has more than one entry or exit records, then the security personnel can come to know if the driver of the car has changed on the basis of the side and front view from the skimmed video.
- The number of cars that entered is also displayed for the user given query.
- The querying can also be done based on the basis of number plate to get the skimmed video of the entry and exit videos which have the car with that number plate.
- We can get information about any car that was present in the parking lot like its structure, color and if there is a dent as the skimmed video will provide with this information.
- User can come to know if the same car has more than one record on the basis of message box displayed.
- The user can get skimmed videos of the car on the basis of structure (hatch/sedan). This is useful when concerned people do not know the number plate of the car but know what shape it had.

B. Scope

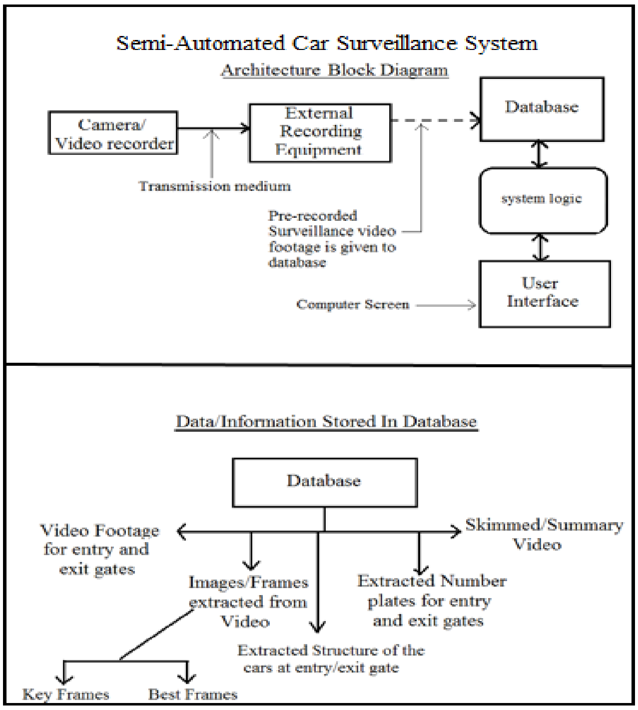
Following are the constraints of the proposed system:

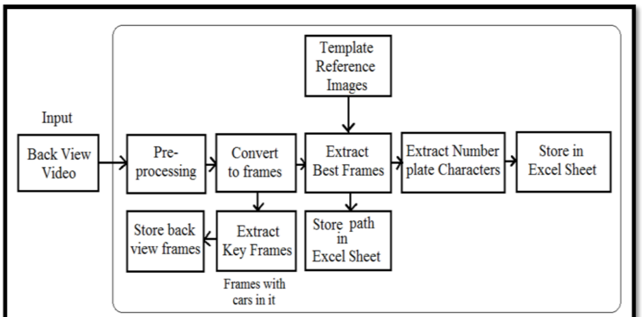
- Pre-recorded video is fetched to the proposed system
- Ideal lighting conditions by using constant light source throughout the day and night, preferably in an enclosed parking space.
- The cameras or the video capturing equipment must be placed in such a way that the light beam rays

incident on the camera are at right angles to the car's surface which is being captured.

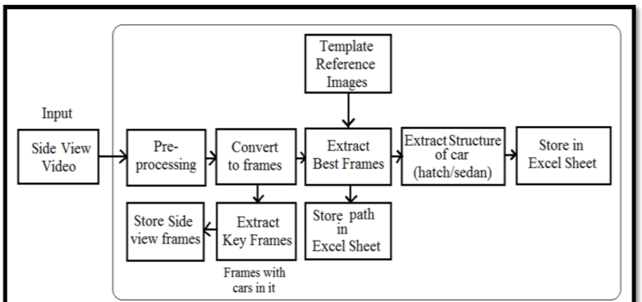
- Three video cameras (DSLR) are placed at different positions such that they record the front, back and side view of the cars.
- The car must move at around 1km/hr to 10km/hr so that the frames do not record motion blur at high speeds.
- The number plate character font we are considering is Arial and its slight variations.
- The number plate should have all the characters in one straight line and the number plate shouldn't be skewed.

C. Architecture





Logic to Store Number Plate Characters from Back View Video Frames



Logic to Extract and Store Structure of Cars from Side View Video Frames

D. Description

A podium parking having an enclosed area was chosen for capturing the videos. Three cameras were placed at both the entry and exit gates respectively. One camera was placed to shoot the back view of the car, 2nd camera shot the front view of the car and the 3rd camera was capturing the side view of the cars at the gates. All the cameras were placed at right angles to the surface of the cars being captured in the video. The videos were then transferred to the storage area i.e. the hard disk (minimum requirement 1TB for surveillance footage of 30 days/1 month). The extension of the videos is .MOV and they are converted to .mp4 to work with. The computer system that we have used to carry out the processing has a CORE i7 processor and 8GB RAM.

The surveillance footage at the entry was processed to determine the cars that are entering and their number plates were stored inside the database using image processing from the back view frames. The frames of the side view of the cars were used to detect the structure of the cars (The two categories we are considering are hatch and sedan). The surveillance video of the cars leaving was also processed similarly.

Processing of Videos to extract Frames, Key frames, Best Frames, Number Plate, and Structure has been achieved using **Matlab** and XML.

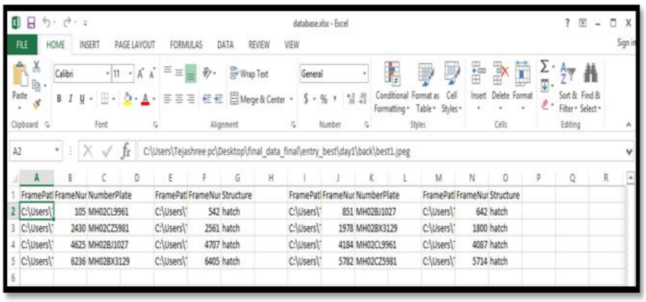
This system is user friendly as it has an interactive GUI. The user will be notified in case of an anomaly or an unusual event that occurs in the surveillance footage on the basis of their query using message boxes. The User Interface has been implemented using **GUIDE** (Matlab's GUI development environment).

E. Hardware Requirements

1) Camera/Video Recorder:
   - Nikon DSLR D3100 – Side View
   - Nikon DSLR D3200 – Front View
   - Canon 700D – Back View
2) Intel Core i3 processor or higher (recommended i7 processor for quick processing and efficiency)
3) Minimum 2GB RAM (recommended 16GB RAM)
4) Minimum Hard Disk Space is 1TB.

F. Software Requirements

1) Operating System : Windows XP, 7, 8 or higher
2) Extensible Mark-up Language
3) MATLAB
   - Matlab R2013a
   - Matlab GUIDE
4) Database: Microsoft Excel



III. METHODOLOGY

The methodology and algorithms used for processing the video data to extract the features and to populate the system's database are as given below.

The videos that are processed are:

- Entry Back View
- Entry Front View
- Entry Side View
- Exit Back View
- Exit Front View
- Exit Side View

Hence, there are 6 videos for each day in a month.

The videos are processed to **extract the frames** that have the cars entering or exiting the parking-area.

A. Implementation Steps

1) Frame Extraction Steps:

- The input video data is read.
- Calculate the total number of frames.
- Divide the video into frames.
- Extract each frame and store it into the database. (In our implementation, this database is a specified folder)
- Stop.

The frames that consist the cars are chosen as key frames to make the general skimmed for the complete day. We have used **the key frame based extraction algorithm to skim the video** into a shorter one to highlight only that part of the video when the car is in the video frames (as there can be many frames without cars and the parking lot may not have cars entering and/or exiting continuously).

2) Key-Frame Extraction

- The input video data is read.
- Refer to the database and compare the consecutive frames.
- Convert the images into grayscale images.
- Compare the images based on correlation factor.

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \bar{A}\right)^2\right)\left(\sum_m \sum_n \left(B_{mn} - \bar{B}\right)^2\right)}}$$

where $\bar{A}$ = mean2(A), and $\bar{B}$ = mean2(B).

Here A and B are the images you are comparing, whereas the subscript indices m and n refer to the **pixel location** in the image. Basically to compute, for every pixel location in both images, the difference between the intensity value at that pixel and the mean intensity of the whole image, denoted as a letter with a straight-line over it.

- Now depending on the value returned on comparison of the consecutive frames, set a threshold value for the extraction of key-frames.
- Frames falling in the range of the threshold value are set as key-frames.
- Extract each key-frame and store it into the database. (In our implementation, this database is a specified folder)
- Stop.

3) Video Skimming

- Create a video file in the database and set attributes like frames per second for the video depending on your requirements.
- Extract the key-frames from the database.
- Include all the key-frames and the frames describing the car into the video file.
- Stop.

4) Best-Frame Extraction

The frames that describe the structure of the car best and the number plate is clearly visible are taken as the **best frames** for side view and back view. The **best frames are binarized** i.e. converted to grey scale and then a threshold value is set to collapse the rest of the background so that the number plate region is visible. Hence the number plate can be detected. In the event of the car and number plate being of the same color, we use edge and boundary detection to identify the number plate region and then crop that image portion to extract the number plate.

5) Number-Plate Extraction

Once, **Number plate is detected**, characters are segmented for number plates having 10 characters based on the spacing between the characters. The **characters are recognized** using character identification by template matching and stored in the database.

6) Structure Extraction

Car detection and Structure classification is carried out. We have considered the side view of the cars to identify them and classify them as either hatch back or sedan. The first step is identifying the car inside the frame which is done using Viola-Jones algorithm and Adaptive Boosting. The second step is to classify the cars which is done by using Point Feature Extraction using SURF.

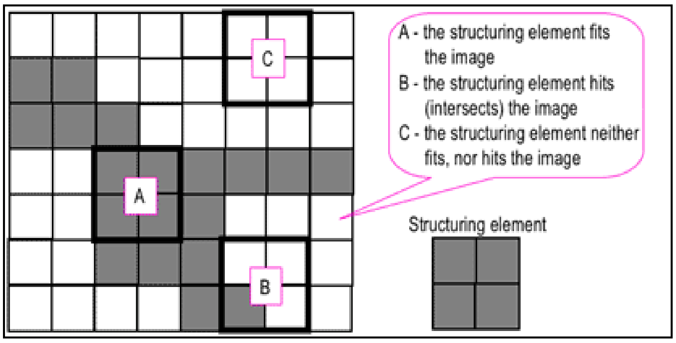
B. Algorithms Used

1) Image Binarization

- Convert RGB to Grey Scale
- Perform Automatic thresholding
- Perform Median Filtering to remove noise

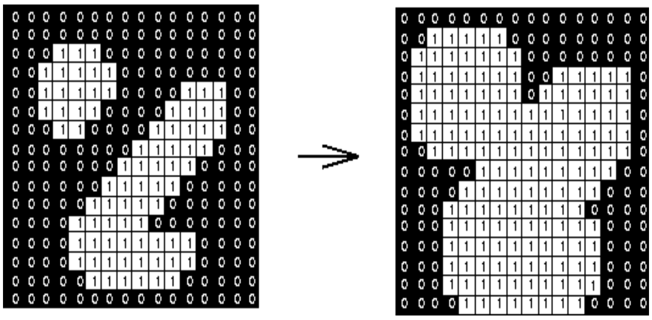
2) Morphological Edge Detection for Number plate localization

- Perform morphological processing to get structural element and remove other insignificant structures.



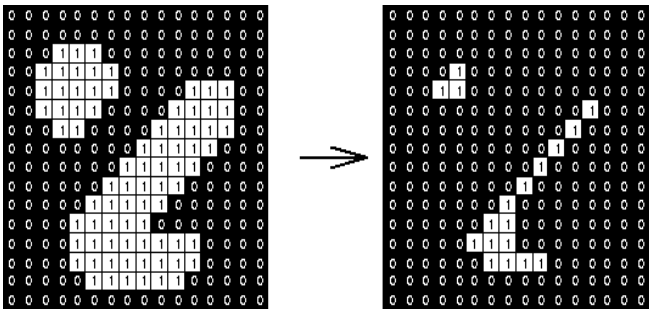
Structuring Element: probe an image with a small shape or template

- Dilate the greyscale image with structural element



Dilation has the opposite effect to erosion -- it adds a layer of pixels to both the inner and outer boundaries of regions.

- Erode this image with structural element



The **erosion** of a binary image f by a structuring element s (denoted f s) produces a new binary image g = f s with ones in all locations (x,y) of a structuring element's origin at which that structuring element s fits the input image f, i.e. g(x,y) = 1 if s fits f and 0 otherwise, repeating for all pixel coordinates (x,y). **Erosion** with small square structuring elements shrinks an image by stripping away a layer of pixels from both the inner and outer boundaries of regions.

- Morphological Gradient for edges enhancement.

3) Morphological Processing for Thinning

- Character Isolation and segmentation is done using Erosion.

4) Character Segmentation

- Crop the characters based on the top and bottom boundary coordinates.
- Characters can also be segmented based on the width and height of template.
- The y-axis length is considered for checking the consistency of the letters.
- Bounding boxes are created by scanning from top to bottom by selecting those elements that have the same vertical length and the elements are separated on the basis of their distance from each other. Thus character segmentation is achieved.

5) Template Matching for Character Recognition

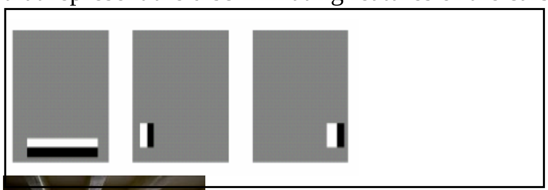
- Pixel values of template characters (A-Z, 0-9) are stored in vector
- Recognized characters are normalized by the template size
- Match the recognized characters with all templates and calculate their similarity
- The best match will be chosen as the result
- The result is converted to text and stored in the database.

6) Viola Jones

The algorithm has mainly 4 stages:

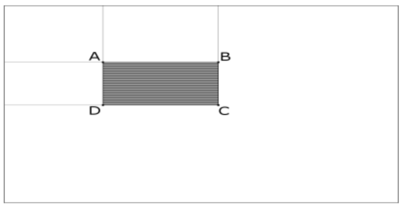
- Haar Features Selection

  The Haar Features were constructed using the knowledge of common characteristics of vehicles such as: the head light and tail light with the wheels below it, and the shadow of the vehicle below it. These are the discriminating features to identify the vehicle. The Haar Features are rectangular filters that represent the discriminating features of the cars.



- Creating Integral Image

  Integral Images are the ones created when the Haar feature window moves over the video frames and detects the discriminating features.



One of the contributions of Viola and Jones was to use summed area tables, which they called integral images. Integral images can be defined as two-dimensional lookup tables in the form of a matrix with the same size of the original image. Each element of the integral image contains the sum of all pixels located on the up-left region of the original image (in relation to the element's position). This allows to compute sum of rectangular areas in the image, at any position or scale, using only four lookups:

$$\text{sum} = I(C) + I(A) - I(B) - I(D).$$

Where A, B, C, D belong to Integral Image I as shown in the above figure.

- Adaboost Training algorithm

It is the Adaptive Boosting Training Algorithm which creates an XML file for the structure of positive samples that are passed to the training network. On the basis of the negative samples, the network classifies them as non-vehicle aspects in the image.

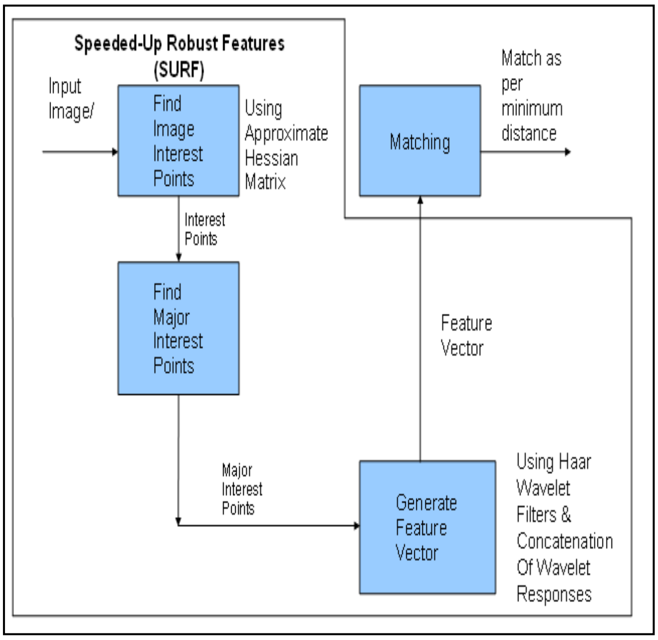
- Cascaded Classifiers

  On the basis of the Adaptive Algorithm, the test images are classified using Cascaded Classifiers. If they have cars in them, then the car is detected and the region of interest is highlighted using a bounding box.

7) Point Feature Extraction using SURF

   SURF: Speed-Up Robust Features

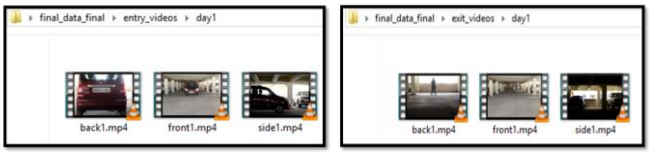
It identifies the object inside a scene by taking several reference object images to classify them as either hatch or sedan.



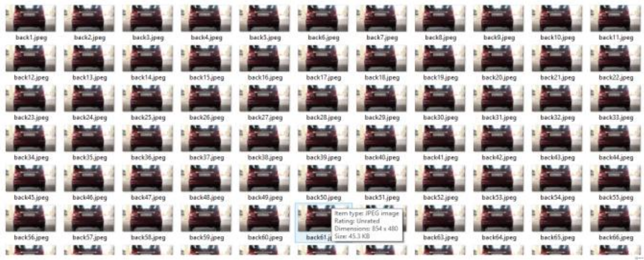
- Read the reference images.
- Read the target best frame side view image.
- Detect Feature points in the reference images and the best frame side view image.
- Visualize the strongest feature points found in the reference images.
- Visualize the strongest feature points in the best frame side view image.
- Extract Feature Descriptors at the interest points in both images.
- Match the features using the descriptors between the reference and best frame images one at a time.
- Get the putatively matched points.
- Locate the structure in the scene using putatively matched points.
- Display detected structure using bounding polygon.
- If the structure has mapped to hatch reference images, then it is a hatch back car else if the structure was mapped to sedan reference images then it is a sedan, else, the car structure wasn't detected.

IV. SYSTEM AND DATA VISUALIZATION

A. Dataset



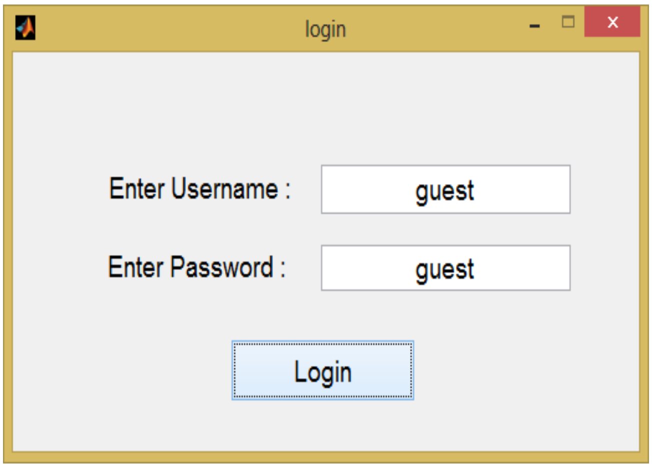
There are 3 videos in each entry and exit folder for every day.



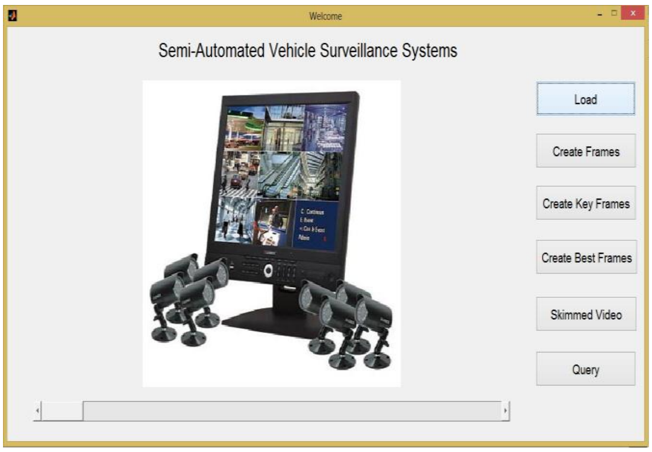
Above are the back view key frames.

These are the side view best frames.

B. User Interface

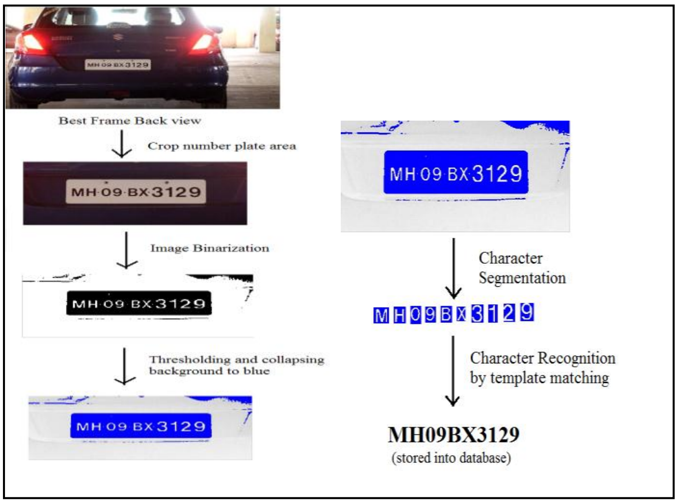


Above is the Login Window to the system.



The welcome window lets the users choose between various operations that they can perform.

C. *Experimental Results*



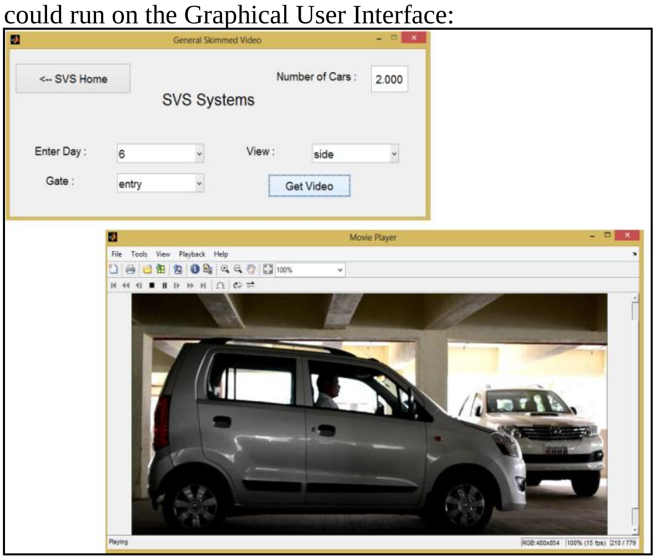
Number-plate detction, character segmentation and recognization

Structure Recognition

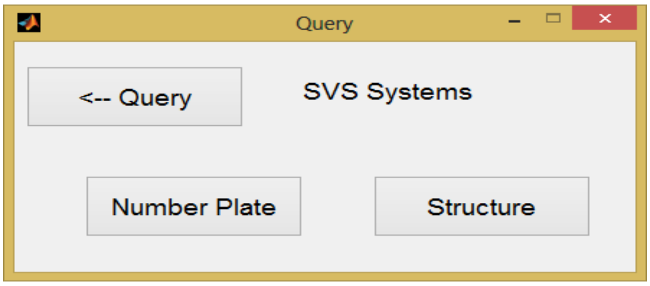
Following are the results of running the previously listed algorithms in each step on our dataset which contains 4 hatch-back cars:

Step	**Car- View**	**Accuracy percentage**
Number Plate Detection	Back	100%
Number Plate Character Recognition	Back	100%
Structure Detection	Side	100%

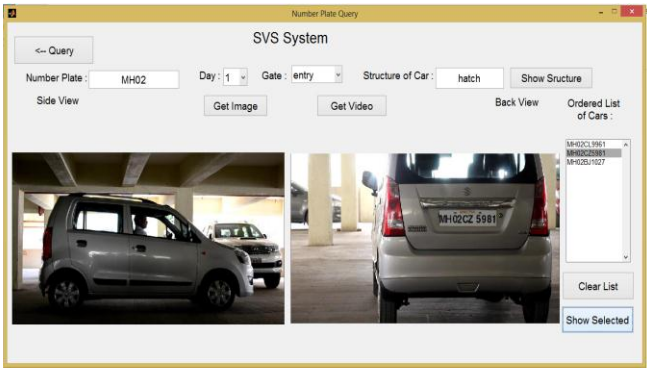
Following are the considered test case queries that the users could run on the Graphical User Interface:



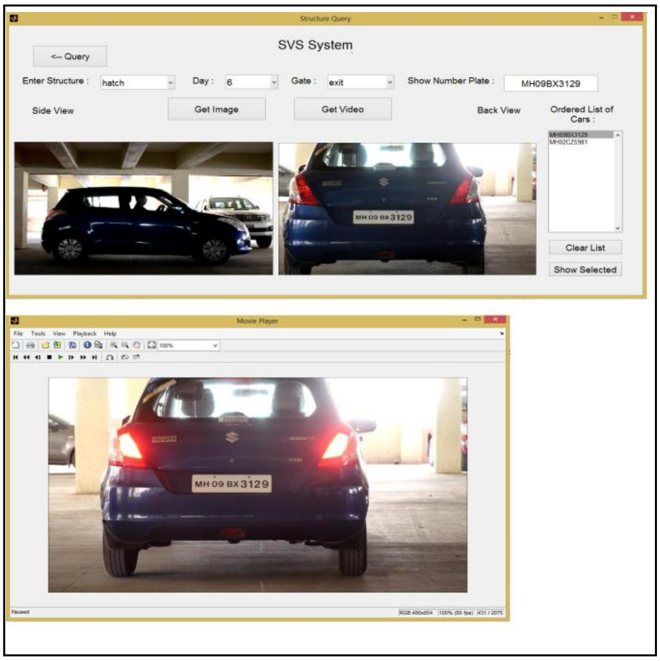
On selecting the day, view and gate, and clicking on the 'Get Video' button in the 'General Skimmed Video' Window, the skimmed video for all cars will be displayed and the number of cars that entered the parking area will be shown on the upper right corner.



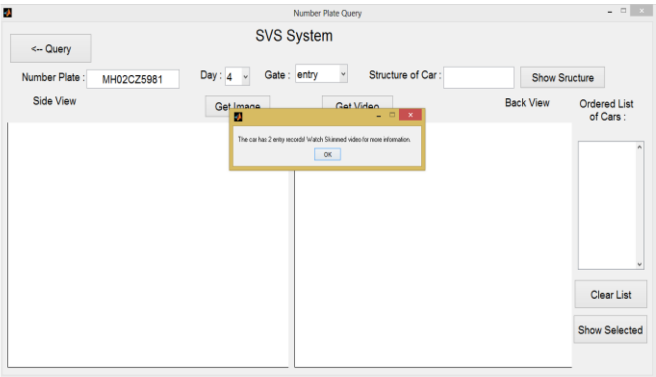
If the user clicks on 'Number Plate' Button, the user is directed to the Number Plate Query Window. If the user clicks on 'Structure' Button, the user is directed to the Structure Query Window.

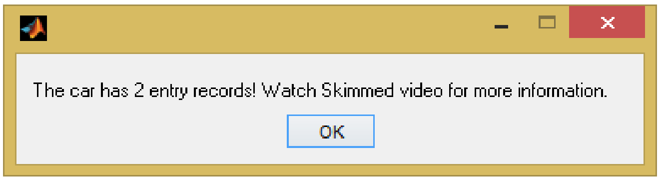


Entering the number, selecting all the remaining choices and clicking on the 'Get Video' button gives skimmed video. Or the user can get the side and back view images of the car if there is only one car pertaining to the query. If there are more than one, then the list of number plates is shown in the list box.



Select between Hatch and Sedan and select the other appropriate choices to get images, skimmed videos and number-plates of the desired cars on click of the buttons.





When there are records that have more than one cars with the same profile, then a message box is displayed indicating an anomaly. Further investigation can be done using appropriate queries given to the system.

V. CONCLUSION

Hence, using Video and Image Processing Technique, it is possible to make an intelligent system that will be able to extract features from the frames of the video and automatically identify required content within the image. It was found that the system was able to achieve the primary goal of Number Plate Extraction and Car Structure Recognition for the scope of this paper.

VI. FUTURE WORK

Further research must be done to help make the system more reliable in terms of character recognition for various fonts and image processing in all types of lighting conditions. Structure detection can be made more specific by exploring neural networks also. Color is another important aspect which can be used to know security related aspects in a surveillance video. The work on extracting color has been chosen as the future scope and incorporating this feature to the system will make it more efficient and user friendly also.

REFERENCES

[1]"Automatic car number plate detection using morphological image processing", Marcin IWANOWSKI, Warsaw University of Technology, Institute of Control and Industrial Electronics, EC Joint Research Centre, Institute of Environment and Sustainability.

[2]"Vehicular Number Plate Recognition Using Edge Detection and Characteristic Analysis of National Number Plates", Bharat Raju Dandu, Abhinav Chopra, Amity University.

[3]"Morphology Based Approach to Recognize Number Plates in India", Phalgun Pandya, Mandeep Singh

[4]" Vehicle License Plate Detection Method Based on Sliding Concentric Windows and Histogram", Kaushik Deb, Hyun-Uk Chae and Kang-Hyun Jo, Graduate School of Electrical Engineering and Information Systems, University of Ulsan, Ulsan, Korea.

[5]"Automatic Number Plate Recognition System (ANPR): A Survey", Chirag Patel, Smt. Chandaben Mohanbhai Patel Institute of Computer Applications (CMPICA), Charotar University of Science and Technology(CHARUSAT); Dipti Shah, PhD. , G. H. Patel P.G. Department of Computer Science, S. P. University; Atul Patel, PhD., Smt. Chandaben Mohanbhai Patel Institute of Computer Applications (CMPICA), Charotar University of Science and Technology(CHARUSAT).

[6]"An Efficient Technique to Recognize License Plate using Morphological Edge Detection and Character Matching Algorithm", Suprokash Dey, Faculty of Engineering and Technology, Jadavpur University, Kolkata, India; Amitava Choudhury, Indian Institute of Engineering Science and Technology, Howrah, India; Joydeep Mukherjee, Faculty of Engineering and Technology, Jadavpur University, Kolkata, India.

[7]"CONTENT BASED VIDEO RETRIEVAL SYSTEMS", by B V Patel1 and B B Meshram, Shah & Anchor Kutchhi Polytechnic, Chembur, Mumbai, INDIA and Computer Technology Department, Veermata Jijabai Technological Institute, Matunga, Mumbai, INDIA.

[8]"Applications of Video Content Analysis And Retrieval", by Nevenka Dimitrova, Philips Research; Hong-Jiang Zhang, Microsoft Research; Behzad Shahraray, AT&T Labs Research; Ibrahim Sezan, Sharp Laboratories of America; Thomas Huang, University of Illinois at Urbana–Champaign; Avideh Zakhor, University of California at Berkeley

[9]"CBIVR: Content-Based Image and Video Retrieval", Prepared by Stan Sclaroff

[10] "Content based Video Retrieval, Classification and Summarization: The State-of-the-Art and the Future", by Xiang Ma, Xu Chen, Ashfaq Khokhar and Dan Schonfeld

[11] "Content-Based Image and Video Retrieval", Dagstuhl Seminar 02021, Schloss Dagstuhl, Wadern, Germany, 6-11 January, 2002. http://www.dagstuhl.de/DATA/Reports/02021.

[12] "Content-based video retrieval", by M. Petkovic, Centre for Telematics and Information Technology, University of Twente.

[13] "Trajectory-Based Video Retrieval Using Dirichlet Process Mixture Models", Xi Li, Weiming Hu, Zhongfei Zhang, Xiaoqin Zhang, Guan Luo (2008), 19th British Machine Vision Conference (BMVC).

[14] "Content Based Video Retrieval: A Database Perspective", By Milan Petkovic.

[15] "Content Based Image And Video Retrieval", By Oge Marques.

[16] "CONTENT BASED VIDEO RETRIEVAL", by Shweta Ghodeshwar, Department of electrical Engineerin, VJTI Mumbai and B.B.Meshram, Department of computer technology, VJTI Mumbai.

[17] "2D Geometric Shape and Color Recognition using Digital image Processing", Sanket Rege, Rajendra Memane, Mihir Phatak, Parag Agarwal.

[18] "Boosted Classifier for Car Detection ," David C Lee, ECE Department, CMU, USA

Tejashree Gharat, Bachelor of Engineering in Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, University of Mumbai.

Tejasvi Dhodi, Bachelor of Engineering in Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, University of Mumbai