A Review on Text Recognition and Traffic Sign Detection

Vaibhav Y.Kukde[a], Neehal B. Jiwane[b], Ashish B. Deharkar[c]*

[a]Student, Shri Sai College of Engineering & Technology, Bhadrawati, Chandrapur 442902, India
[b]Assistant Professor,Shri Sai College of Engineering & Technology, Bhadrawati, Chandrapur 442902, India
[c]Assistant Professor,,' Shri Sai College of Engineering & Technology, Bhadrawati, Chandrapur 442902, India

A B S T R A C T

In this review paper detection and discovery of text in signs of traffic Scene building from Canny edge detector with Hough transform and HOG descriptors is used to summary search regions inside the image, within which traffic sign candidates or models are then found. Maximally stable extremal regions with HSV or HSI color thresholding are used to find a tall variation of candidates, and then they're decrease by put on restraints supported by temporal and operational information. The identification stage interprets the text contained within detected candidate regions. Maximally stable extremal regions are Individual text characters detected and are arranged into lines, before being taken using Support vector machine and optical character recognition. Immensely improved Recognition accuracy by through the temporal fusion of text results across successive frames.

Keywords: Maximally stable extremal region, Hue Saturation Value or Intensity color, Optical Character Recognition and support vector machine

1. Introduction

The instinctive recognition and detection of traffic signs is a problematic in now a days, with a variability of important application areas, together with independent vehicles and progressive driver assistance systems, road surveying. This system includes two main stages: recognition and detection. The detection stage feats information from shaped based detection i.e., the size locality and color based detection of the road in   frame, to see the regions within the scene for traffic text signs is to be searched. These regions are defined once the vanishing point (VP) of the scene is known and hence, determined the ground plane by canny edge detector with Hough transform and HOG descriptor after that scene search regions are introduced and then color based detection is implemented by using a combination of maximally stable extremal regions, hue, saturation, and value or moderation color thresholding.

By matching these regions through successive frames, sequential data is used to additionally removing detected regions, based on the gesture of regions with respect to the camera and also the construction of the scene.

Once a probable traffic sign has been located, successive stage of the algorithm makes an attempt to recognize text among the region before estimate perspective transform is applied to the region, in order to get shaped from unshaped region. Candidate working for text characters are then sorted into potential text lines and located among the region, before being taken using off-the-shelf optical character recognition package or supportvector machine.
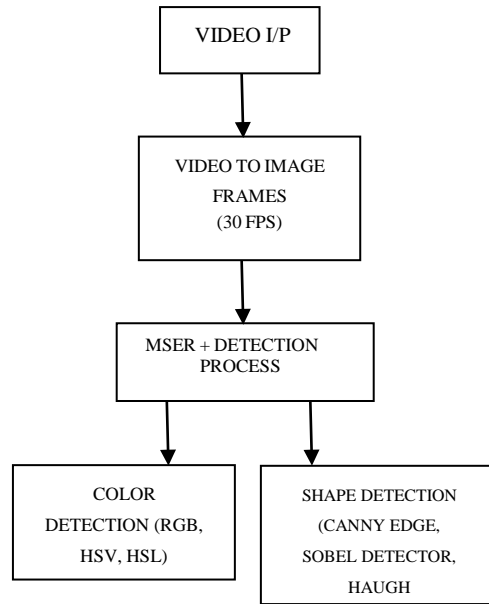
To increase the accurateness of recognition, optical character recognition results from many frames are combined together by similar individual words through frames using a weighted histogram and then detection and recognition progration are described below

1. Detection

* Corresponding author. Tel.: +919767083411

E-mail address: kukdevaibhav9@gmail.com; neehaljiwane@gmail.com

In the detection stage, two approaches occur. Detection based on shape criteria of traffic signs and color. The combination of the two progressions is also feasible. Basic shape detection and color detection are shown in figure1

```
          ┌──────────────┐
          │  VIDEO I/P   │
          └──────┬───────┘
                 │
                 ▼
          ┌──────────────────┐
          │  VIDEO TO IMAGE  │
          │     FRAMES       │
          │    (30 FPS)      │
          └────────┬─────────┘
                   │
                   ▼
          ┌──────────────────┐
          │  MSER + DETECTION│
          │     PROCESS      │
          └───┬──────────┬───┘
              │          │
              ▼          ▼
    ┌──────────────┐  ┌──────────────────┐
    │    COLOR     │  │ SHAPE DETECTION  │
    │ DETECTION (RGB,│ │ (CANNY EDGE,     │
    │  HSV, HSL)   │  │ SOBEL DETECTOR,  │
    │              │  │   HAUGH          │
    └──────────────┘  └──────────────────┘
```

A. Maximally stable extremal region and Hue Saturation Value Thresholding

The algorithm includes detecting the candidates for text- based traffic signs among the bounded scene search regions it taken from our past or previous work in the detection of symbol and text based traffic signs candidates using each Maximally stable extremal region and Hue Saturation Value color thresholding.

These two types of detector are used, to reach a color detection make sure that all possible traffic signs are detected in all situations. Maximally stable extremal region are outlined to be regions that preserve their shape approximately through several image threshold levels. This region detector is robust to illumination and contrast differences and detect high- contrast regions, that build it approximate for the detection of traffic signs. Further traffic text sign candidates are detected using Hue Saturation Value color thresholding some values of Hue Saturation Value colorthresholding are shown in table 1.

Each frame is first remodeled into the HSV color space, before a threshold is applied to each hue and saturation channels. The value channel is neglected to help the system remain invariant to changes in brightness.

The Maximally stable extremal region detection shown in figure2 Region detection should be repeatable and stable, and capable to discriminate between the regions

COLOR	HUE RANGE	MIN SAT	MAX SAT
BROWN	12°-52°	50%	100%
GREEN	136°-176°	20%	100%
BLUE	184°-224°	24%	100%

Table.1.value used for HSV thresholding

The algorithm

i. Start from an area intensity extremum point
ii. go in each direction until the point of maxima of few function. The curve connecting points is the regionboundary
iii. compute geometric moments of orders up to two for this region
iv. Replace the region with ellipse

Processing the MSER extraction implements the subsequent steps:

i. Sweep threshold of intensity from black to white, performing an easy light thresholding of the image

ii. Extract connected elements

iii. Notice a threshold when an extremal region is "Maximally Stable".

iv. Approximate an area with Associate in the form of ellipse (this step is optional) and Keep those regions descriptors as options

B. Hsi Thresholding

Candidate objects are choosen via thresholding. Thresholding refers to the procedure that makes a binary image; pixels with illumination values higher than a predefined threshold are assigned value 1, and all the others are set to zero. Thresholding are controlled in HSI (Hue – Saturation – Intensity) color space, as it is more strong to illumination changes than RGB. Only Hue Saturation channels are used, as these components encode color info. possible values for Hue element vary between 0°-360° and for the Saturation element between 0- 255. some values of HSI thresholding are give in table 2

COLOR	HUE RANGE	SATURATION
BLUE	180<H<290	S>102
RED	H<22(or)H>306	S>38
YELLOW	18<H<72	0.5<S<204

Table.2. values used for HSI thresholding

C. Hough Transform And Canny EdgeDetector

After color based detection, elimination of false positives and shape based detection is performed so as to separate overlapping traffic signs. shape based detection takes advantage of canny edge detector, Hough transform (line and circular).circular Hough transform that is the algorithm that identifies circular curves. The circular Hough transform is to be applied, an edge image, acquired via Canny detector, is crucial. Afterwards, the presence of circles is indicated through a voting procedure carried out in the parameter domain. the number of dimensions of the parameter space equals to the number of parameters needed to completely outline the curve. As a circle is mathematically expressed through the equation (1), voting procedure includes the position x_o, y_o of the center of the circle and its radius r.

$$(x - x_o)^2 + (y - y_o)^2 = r^2 \text{------------------------} (1)$$

These three parameters from the accumulator array and combinations with the highest values of votes are more possible to represent circles. Circular Hough transform is applied to ROIs (Region of interests) with aspect ratio less than 0.7, as these ROIs are more anticipated to depict overlapping road signs. If a circle is detected, its bounding box is derived and the area between the new and the original bounding box forms a new region of interest. This new ROI is accepted only if the quantitative relation between the vertical sides of the two new bounding boxes (BB of the circle and remaining BB) is between 0.7 and 1.3.

D. Hog Descriptors

For the recognition stage, regions of interest are described using histogram of oriented Gradients (HOG) proposed by Triggs and Dalal (2005). HOG descriptors have firstly been applied for pedestrian detection but till then, they have been widely used for object recognition, as they're sturdy to scale and illumination changes. in order to extract HOG descriptors, an image is divided into blocks, which are formed by overlapping cells. every cell consists of non-overlapping pixels and for every cell a local 1-D histogram of edges orientation is derived. each pixel contributes to the formulation of the histogram by the magnitude and also the orientation of its gradient. Orientation angles are measure into bins and for each pixel a vote is assigned to the suitable bin to which the orientation value belongs and this vote is assigned by the magnitude. the number of bins is changeable and also the range of orientation is between 0°-180° for unsigned gradient and 0°-360° for signed gradient. The local histograms are accumulated over blocks in order to attain illumination invariance and are then concatenated to createthe descriptor.

2.Recognition

The recognition procedure is to assign every region of interest to the category that it belongs. Also, fps (false positive) that have been detected as candidates depicting a road sign and that are eliminated. Recognition shall be performed using traditional template matching from the field of machine learning, like Optical character recognition (OCR) and Support Vector Machines
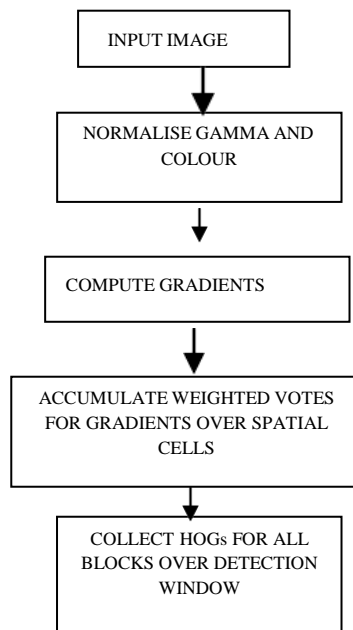
```
            ┌─────────────────────┐
            │     INPUT IMAGE     │
            └─────────────────────┘
                       │
                       ▼
         ┌─────────────────────────┐
         │  NORMALISE GAMMA AND    │
         │         COLOUR          │
         └─────────────────────────┘
                       │
                       ▼
         ┌─────────────────────────┐
         │    COMPUTE GRADIENTS    │
         └─────────────────────────┘
                       │
                       ▼
         ┌─────────────────────────┐
         │ ACCUMULATE WEIGHTED VOTES│
         │ FOR GRADIENTS OVER SPATIAL│
         │          CELLS          │
         └─────────────────────────┘
                       │
                       ▼
         ┌─────────────────────────┐
         │  COLLECT HOGs FOR ALL   │
         │  BLOCKS OVER DETECTION  │
         │         WINDOW          │
         └─────────────────────────┘
```

Fig.1. Basic color detection and shape detection

a.   Optical Character Recognition

Optical Character Recognition (OCR) is the electronic or mechanical translation of images of written or typewritten text into machine-editable text. The main aim of this project is to design an expert system which can be best to, OCR that can effectively recognize a specific character of sort format exploitation the Feed Forward approach. OCR could be a field of research in artificial intelligence [5-6], in pattern recognition and additionally in machine vision

Although tutorial analysis within the field that continues, the main focus on OCR has been shifted to implementation of well-tried techniques.

OCR (uses optical techniques like lenses and mirrors) anddigital character recognition (using computer algorithmsand scanners) were originally thought as separate fields. asa result a very few applications survive that use actualityoptical techniques, the OCR term has been broadenedcurrently to incorporate digital image processing as well.this technique will be applicable to recognize any varietyof characters together with uppercase, lowercase alphabetsand numerals. Basic process of OCR is given in figure4 OCR engine process

  i.   The system acquires an image of the scanned paper.
 ii.   The scanned paper is given as input to the OCR engine.
iii.   The portions of the image are matched to shapes by engine that is instructed to recognize.

b.   Support Vector Machine

SVMs (support vector machine) are a bit difficult. in this case, it tend to show a linear SVM and illustrate its behavior on some 2d data. this should be great for aiming to grips with maximizing geometric margins, support vectors, and the optimization perform an maximum separating hyper plane[1]. Data can be generated randomly (uniformly or from separate Gaussians) over the 2d space, and an SVM can be trained to find a separate line. data points can be dragged around with the mouse, and the model (SVM) will retrain in real-time as the point is dragged (observe that dragging non-support vector points will not affect the SVM decision boundary) are shown in figure5

The program needs some implementation of A SVM algorithm. Therefore, you'll have to compel to have one of:

i.   The bioinformatics toolbox and third-party library "libsvm", which includes an SVM train function
ii.   The optimization toolbox, which includes a quad prog function

3. Conclusion

The automatic recognition and detection of text in traffic signs supported by MSERs with HSV or HSL thresholding, HOG descriptor. The search space for traffic signs was reduced exploitation structural information from the scene that assisted in reducing the total number of fps. Perspective rectification and temporal fusion of candidate regions of text were used to improve OCR and SVM results.

References

1. S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez- Moreno, and F. Lopez-Ferreras, "Road-sign detection and recognition based on support vector machines," IEEE Trans. Intell. Transp. Syst., vol. 8, no. 2, pp. 264–278, Jun. 2007

2. J.Greenhalgh and M. Mirmehdi, "Traffic sign recognition using MSER and random forests," in Proc. EUSIPCO, Aug. 2012, pp. 1935–1939.

3. J.Greenhalgh and M. Mirmehdi, "Real-time detection and recognition of road traffic signs," IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1498–1506, Dec. 2012.

4. Mogelmose, M. M. Trivedi, and T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4,pp. 1484–1497, Dec. 2012.

5. Google, Tesseract-OCR 2013, [Online; accessed 8-October-2013]. [Online].\Available:http://code.google.com/p/tesseract-ocr/

6. The Department for Transport, "Working drawings for traffic signs,"London, U.K., Oct. 2013. [Online]. Available:https://www.gov.uk/working-drawings-for-traffic-signs

7. Pratiksha S. Bodhe, Neehal B. Jiwane published paper in International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE) on the topic " Design and Implementation of E-learning System" on Vol. 11, Issue 5, May 2022

8. Viraj Kalambe, Neehal B. Jiwane published paper in International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE) on the topic " College Enterprises and Resources Planning" on Vol. 11, Issue 5, May 2022

9. Sujata Bhalme, Neehal B. Jiwane published paper in International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE) on the topic " Digital voting system using on blockchain" on Vol. 11, Issue 5, May 2022