Road-Sign Detection and Recognition Based on Support Vector Machines

Maldonado-Bascon et al. *et al.*

Presented by

Dara Nyknahad

ECG 789

Outline

- Introduction
- Support Vector Machine (SVM)
- Algorithm
- Results
- Conclusions

Introduction and View

- What?
  - Automatic road-sign detection and recognition system based on support vector machines (SVMs)

- Why?
  - Road signs are important
  - Provide drivers important information
  - Help them to drive more safely and more easily
  - Guiding and warning

Introduction and View

- Input?
  - Video-> .bmp frames
- Our desire?
  - ?

TABLE I
MEANING OF SPANISH TRAFFIC SIGNS ACCORDING TO THE COLOR AND SHAPE

Color	**Shape**	**Meaning**
Red Rim	Circle	Prohibition
Red Rim (Up)	Triangle	Danger
Red Rim (Down)	Triangle	Yield
Red	Octagonal	Stop
Blue	Square	Recommendation
Blue	Circle	Obligation
White	Circle	End of prohibition
Yellow	Circle	End of prohibition (construction)

Introduction and View

- Input?
  - Video-> .bmp frames
- Our desire?
  - Color
  - Shape
  - Content

Detection

TABLE I
MEANING OF SPANISH TRAFFIC SIGNS ACCORDING TO THE COLOR AND SHAPE

Color	**Shape**	**Meaning**
Red Rim	Circle	Prohibition
Red Rim (Up)	Triangle	Danger
Red Rim (Down)	Triangle	Yield
Red	Octagonal	Stop
Blue	Square	Recommendation
Blue	Circle	Obligation
White	Circle	End of prohibition
Yellow	Circle	End of prohibition (construction)

Introduction and View

- Input?
  - Video-> .bmp frames
- Our desire?
  - Color
  - Shape
  - Content

Detection

Recognition and interpretation to knowledge

TABLE I
MEANING OF SPANISH TRAFFIC SIGNS ACCORDING TO THE COLOR AND SHAPE

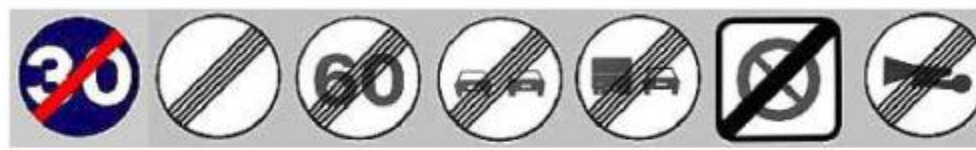
Color	Shape	Meaning
Red Rim	Circle	Prohibition
Red Rim (Up)	Triangle	Danger
Red Rim (Down)	Triangle	Yield
Red	Octagonal	Stop
Blue	Square	Recommendation
Blue	Circle	Obligation
White	Circle	End of prohibition
Yellow	Circle	End of prohibition (construction)

Introduction and View

- For what application?
  - Driver-Assistance Systems
  - Maintenance purposes

Introduction and View

- What are the difficulties?



Introduction and View

- What are the difficulties?

Different Shapes
Different Colors
Different Contents
No More?

Introduction and View

- What are the difficulties?
  - Variable Lighting Conditions
  - Variable Sign Rotation
  - Variable Sign Dimensions
  - Invisibilities (natural like trees, others like vehicles)
  - Large class of Signs
  - Color Combination

Algorithm

- The system consists of 3 stages:
  - 1) Segmentation according to the color of the pixel;
  - 2) Traffic-Sign detection by Shape classification using linear SVM
  - 3) Content Recognition based on Gaussian-kernel SVMs

Algorithm Overview

- 1) Segmentation
- 2) Shape classification
- 3) Content Recognition

The complete process is triggered by color segmentation of the frame, where the system will search objects with similar colors as traffic signs: RED, BLUE, YELLOW, WHITE

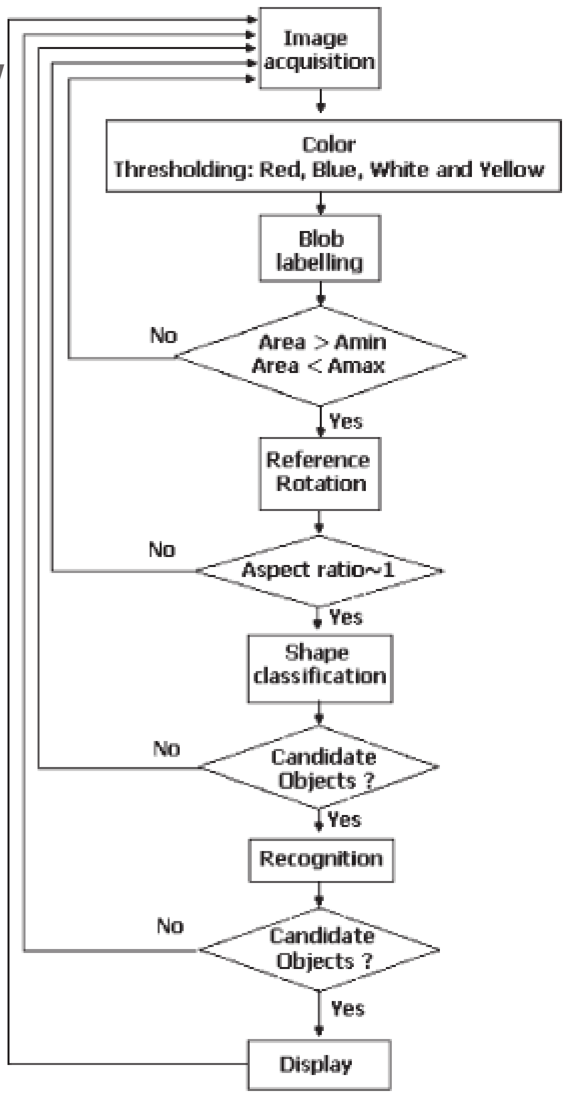


Fig. 2. Algorithm description.

Algorithm

- 1) Segmentation
- 2) Shape classification
- 3) Content Recognition

Rejection of similar object like CAR, BUILDINGS, … happens in few places based on some criteria:

1.Geometcric feature selection

2. Shape classification

3. Recognition of inner area

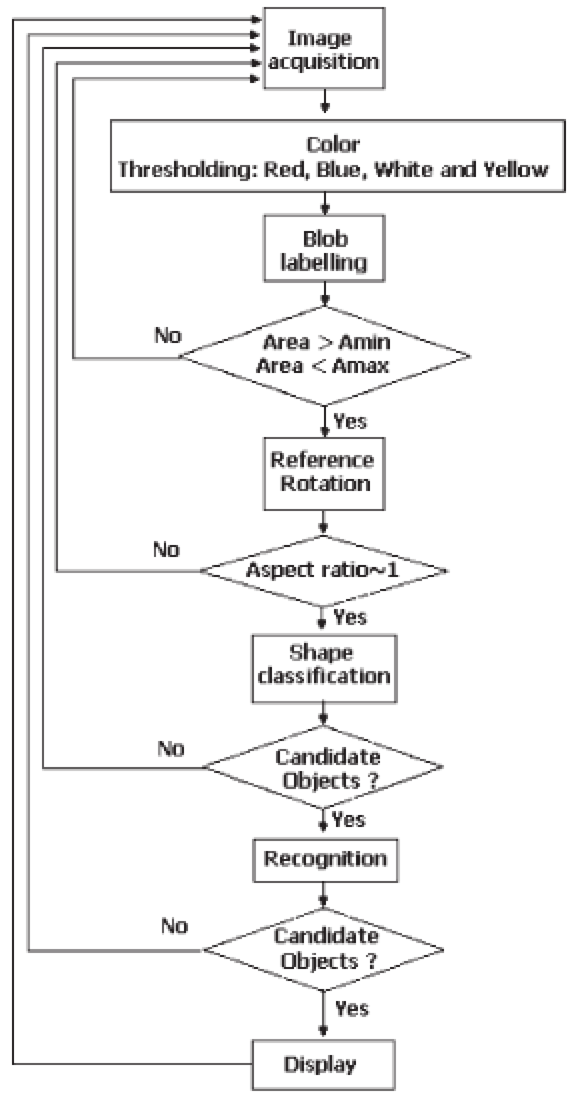


Fig. 2. Algorithm description.

Algorithm

- 1) Segmentation
  - Candidate blobs are extracted from the input image by thresholding using HSI color space for chromatic signs. At the same time, white signs are detected with the help of an achromatic decomposition.
- 2) Shape classification
  - Blobs that are obtained from segmentation are classified in this stage using linear SVMs. According to the color that has been used in the segmentation, only some given shapes are possible. For example, signs that are segmented using the red clues can be circular, triangular, or octagonal.
- 3) Content Recognition
  - The recognition process is based on SVMs with Gaussian kernels. Different SVMs are used for each color and shape classification.

Segmentation

- Different color space used for segmentation and HIS was selected.
- Difficulties in segmentation
  - As we discussed illumination and deterioration, ...

Color Domain

- Why HIS?
  - H and S are having Low variation for objects of interests with similar color.

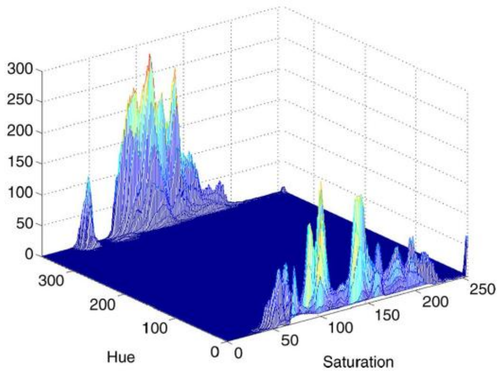


Fig. 3. Hue–saturation histogram for red signs.

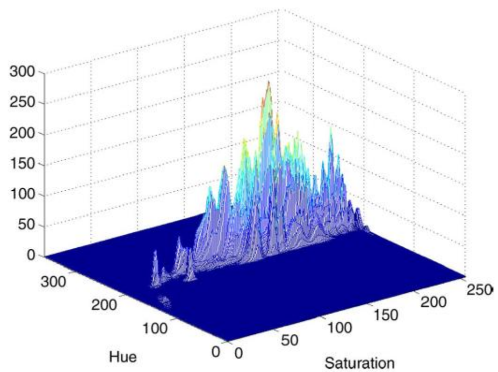


Fig. 4. Hue–saturation histogram for blue signs.

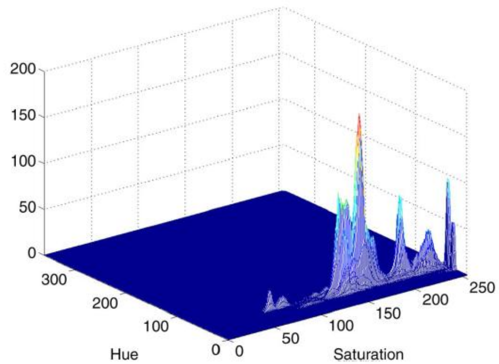


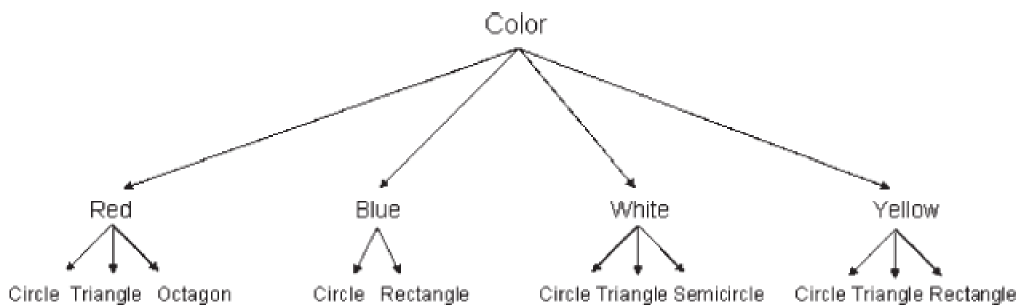
Fig. 5. Hue–saturation histogram for yellow signs.

Color Domain

- Segmentation results ➡ boundary of interest (BI)

Color

Shape



(a) (b) (c)

Fig. 7. Segmentation results.

White Signs

- Unfortunately, the hue and saturation components do not contain enough information to segment white signs.

- The image's achromatic decomposition then helps to detect white

$$f(R,G,B) = \frac{(|R-G| + |G-B| + |B-R|)}{3D} \qquad (1)$$

D is the degree of extraction of an achromatic color,

An *f(R,G,B) of less than 1 represents achromatic* colors, and an *f(R,G,B) of greater than 1 represents chromatic* colors. (NOT GOOD FOR NIGHT)

Segmentation

- All candidate blobs are analyzed in a selection process
- Some of them are discarded according to their size or aspect ratio
- Small blobs are rejected as noise
- Big blobs are rejected as noninterest objects

The result of color segmentation

(Possible Traffic Signs)

will be fed to shape classification

Algorithm

- 1) Segmentation

- 2) **Shape classification**

- 3) Content Recognition

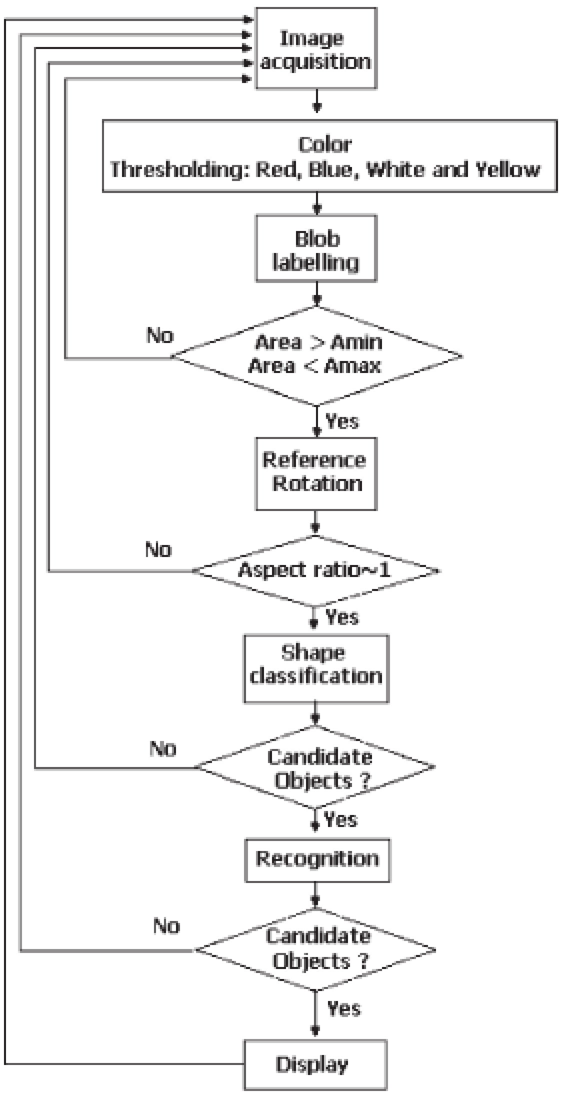


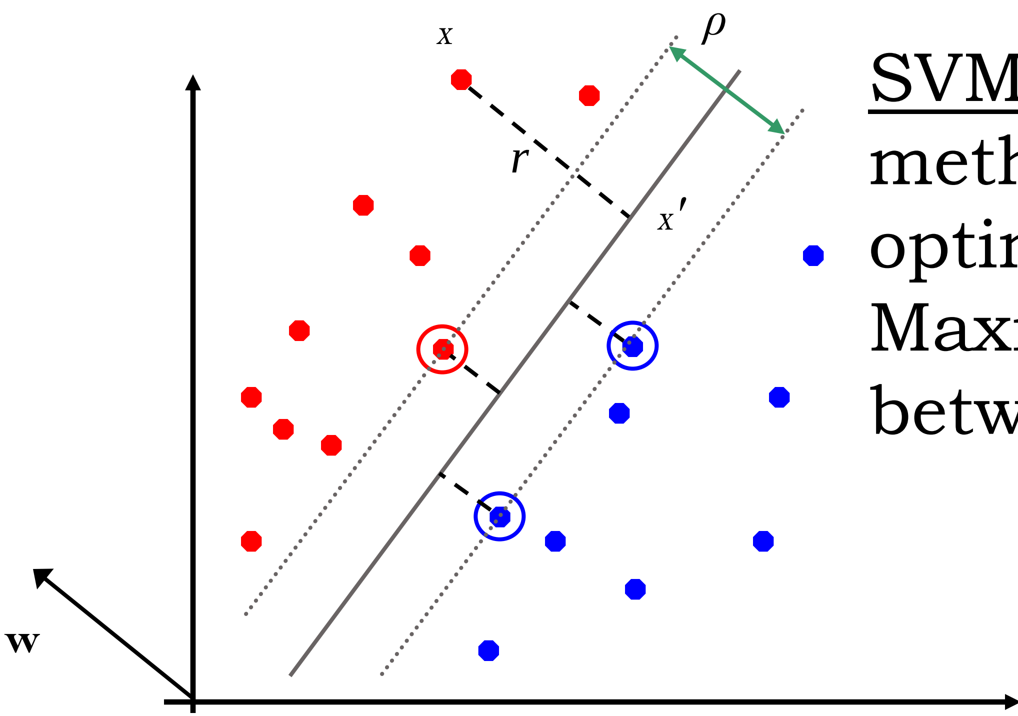
Fig. 2. Algorithm description.

Shape Classification

- The result of Segmentation is fed to Shape Classification
  - The BI are being classified into different shapes.
  - The linear SVM is minimizing the upper bound (instead of error of training data) to minimize structural risk.
  - Although it was mainly used for text classification but they used it for binary classification (0/1)

Support Vector Machine

- A linear classifier $ax + by - c = 0$

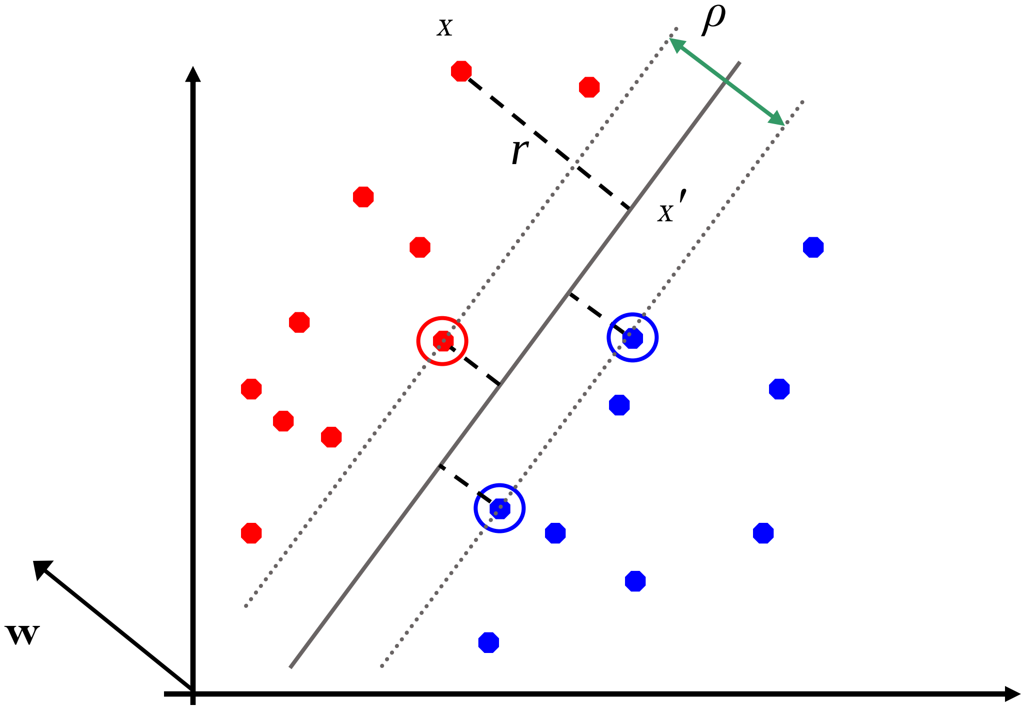


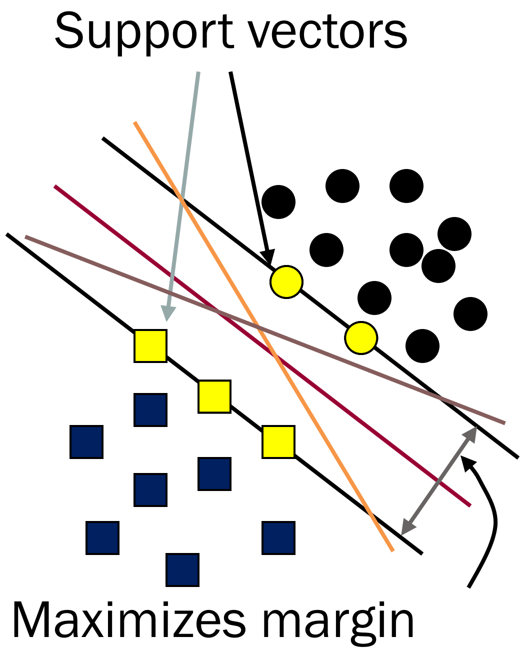
SVM an optimization method that finds an optimal solution. Maximizes the distance between the hyperplanes

Hyperplane H is a set of points {x1, x2,…} that satisfy a linear relation ∑Axi=B

Support Vector Machine

- A linear classifier $ax + by - c = 0$





Support Vector Machine

- SVM maximizes the *boundary* around the separating hyperplane
- SVM is $O(n^2)$
- Usually is used for text classification

Shape Classification

- Two Separable classes yi {-1 or 1}
- Training data set {xi,yi}
- The vector xi are the DtBs (distances from the external edge of the blob to its bounding box.)
- If the hyper plane {w,b} separates the two classes, the points that lie on it satisfy the $x.w^T+b=0$ (w is norm to plane)
- So this leads to satisfying:
  - For all i: $yi(xi.w^T+b)-1>=0$
  - and the margin: 2/||w|| and maximizing it is desired. be achieved by minimizing

$$Lp = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l} \alpha_i y_i \left(\mathbf{x}_i \cdot \mathbf{w}^T + b\right) + \sum_{i=1}^{l} \alpha_i.$$

Shape Classification

- So by using $f(\mathrm{x}) = \mathrm{sgn}(\mathrm{x} \cdot \mathrm{w}^T + b)$

  we determine that a given vector x is belonged to which class.

- Visual example of DtBs (distances from the external edge of the blob to its bounding box.

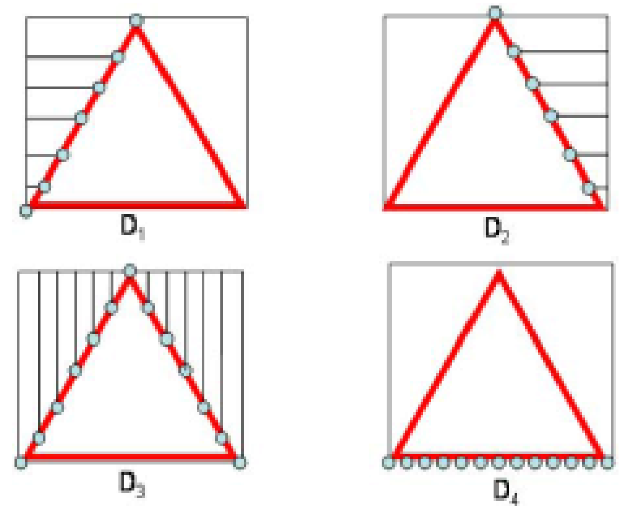


Fig. 9. DtBs for a triangular shape.

Some Notes on Shape

- The position of the candidate blob does not matter
- Invariant to
  - Scale
  - Rotation (2D, 3D)
  - Translation
- Good results with blind spots (occlusions)

Algorithm

- 1) Segmentation
- 2) Shape classification
- **3) Content Recognition**

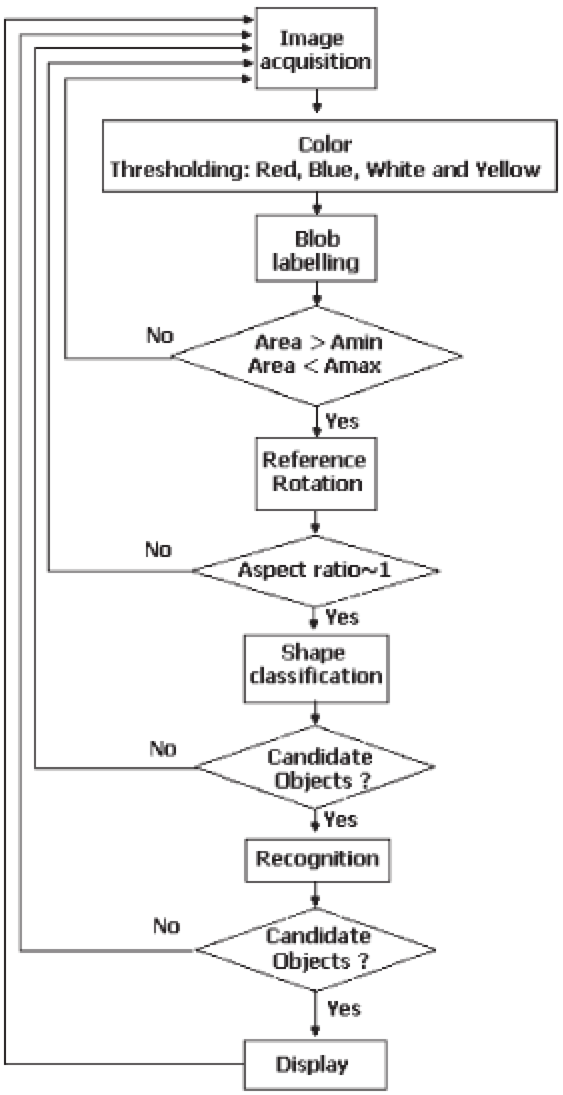


Fig. 2. Algorithm description.

Recognition

- Once the candidate blobs are classified into a shape class, the recognition process is initiated.
- Recognition is implemented by SVMs with Gaussian kernels
- Why? In many cases, the data cannot be separated by a linear function.
- How? A solution is to map the input data into a different space
- The recognition stage input is a block of 31×31 pixels in grayscale image for every candidate blob
- Therefore, the interior of the bounding box is normalized to these dimensions.

Recognition

- Gaussian kernels: $K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}}$
- Decision maker function: $f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{N_s} \alpha_i y_i K(\mathbf{s}_i, \mathbf{x}) + b\right)$
- Average of training samples: [20 100] avg=50 for each class

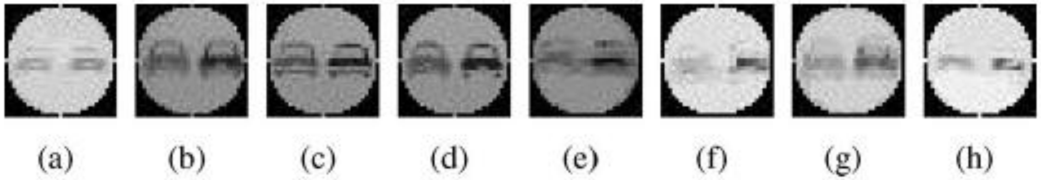


Fig. 17. Positive support vectors for the "No overtake" traffic sign by achromatic segmentation.

Recognition

- Gaussian kernels: $K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}}$

- Decision maker function: $f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{N_s} \alpha_i y_i K(\mathbf{s}_i, \mathbf{x}) + b\right)$

- Average of training samples: [20 100] avg=50 for each class

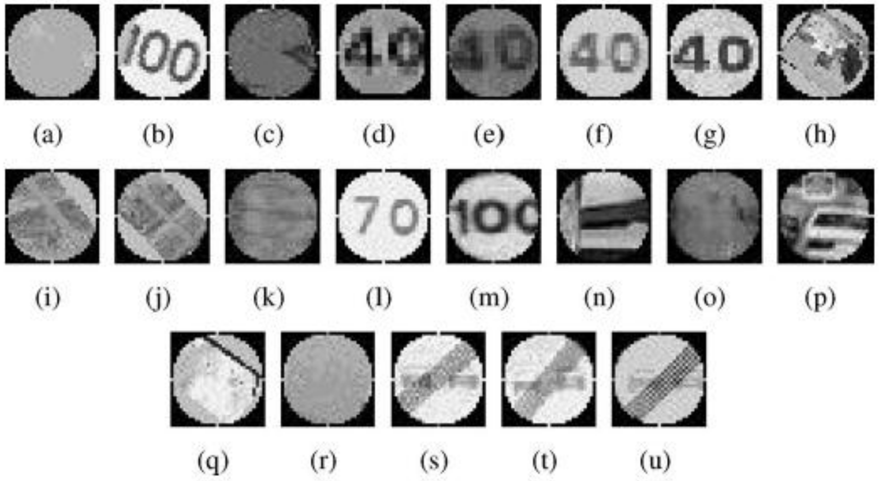


Fig. 18. Negative support vectors for the "No overtake" traffic sign by achromatic segmentation.

Recognition

- There are some values and parameters that should be set like thresholidng values for discarding noise blobs.
- Effect? These parameters could be set to change:
  - False alarm probability
  - Lost probability

There are always some exception cases that could be handles via separate training sets

Summary

- The experiments run at driving at normal speed for 4 km, day and night.
- All the signs have been detected at least twice

TABLE III
SUMMARY OF RESULTS

Number of sequence	1	2	3	4	5
Number of images	749	1774	860	995	798
Number of traffic signs	21	21	20	25	17
Detections of traffic signs	218	237	227	285	127
Noisy candidate blobs	601	985	728	622	434
False alarms	0	3	4	8	7
Confused recognition	4	4	4	2	7
	Sunny	Sunny	Sunny	Rainy	Night

Summary

- An important conclusion from the results is that false alarms do not appear in the same sequence several times, and so they could be rejected by the tracking algorithm.
- The system also works when the signs are not placed perpendicular to the movement of the vehicle (3-D rotations).
- The system is generally able to recognize objects with so many different scales as standard traffic signs.

Summary

- For the case of occlusion

(a) (b) (c)

Fig. 26. Masks of occlusion. Three sizes are used to test the robustness of the system against occlusions with different areas covered. (a)–(c) Masks whose diameters are one half, one third, and one fourth of the major dimension of the bounding box, respectively.

Summary

- For the case of occlusion

(a) (b) (c)

Fig. 27. Displacements followed by a mask for the first orientation.

44.90% 67.85% 93.24%

(Recognition success probabilities)

Summary

- A comprehensive methodology to:
  - Detect traffic signs
  - Recognize traffic signs
- Using video and convert to frames as input
- Considering difficulties (illumination, rotation, occlusion, …)
- Invariant to rotation, scales, displacement
- The minimum 2 frames detection/recognotion was maintained for detection and recognition
- Good results in occlusions

Thank you