Optimization of Road Sign Detection Based on Colour information for OCR

Eric Gambardella and Shan Fu
Centre for Computer Vision and Image Processing
School of Engineering, Cranfield University, United Kingdom

Abstract

Road sign recognition is a driving assistant function in foreseeable future. This function produces not only the detection of road signs but also the interoperation the content of the signs. In order to obtain the most reliable result in a short processing time, one of the strategies is to find the most suitable image of the sign for the optical character recognition (OCR). This paper proposes an automated procedure to detect road signs, and to track them until the image of the best quality is found. We used multi-scale colour clustering procedure to find road signs, apply a linear motion based tracking procedure to increase the robustness of the detection, and propose a statistic measurement combining all the colour "behaviour" within the road sign region to indicate the quality of the road sign image. The proposed procedure is tested on the image sequences obtained from British and French motorways.

Introduction

While driving, driver could sometimes miss road signs due to the lack of concentration. It would be a useful assistant function to recall road signs the driver in reverse chronological order.

Many applications of vision systems concerning road sign detection and recognition have been developed recently.[1-4] The automatic detection and classification of road signs is clearly an emerging research topic in the field of intelligent vehicles. Different techniques have been proposed for road sign detection from a sequence of monochromatic or colour images.

In the method presented in reference,[1] road sign detection and classification consists of three principal steps. In the first step, a colour segmentation algorithm was applied to the image. In the second step, regions of interest were extracted using colour combinations, which were characteristics for traffic signs. Finally, the regions of interest had to be tested by a pictogram classifier.

Picciola *et al.* presented a method for detecting and recognizing road signs in both grey-level and colour images. The method worked in three stages. Firstly, the search region was extracted using some *priori* knowledge or colour clues. Secondly, a shape analysis was applied to find circular and triangular signs in the edge map. Finally, the sign was classified using a cross-correlation with a database of road signs. The detection algorithm was also improved using temporal integration based on Kalman filtering methods.

The system presented in Ref. 3 was a method to perform detection and recognition of triangular and circular road signs. Firstly, the road sign position was found using *a priori* information of shape and colour. Then the road sign was extracted from the closer view image using conventional template matching. The recognition phase consisted of two processes: training and testing based on matching pursuit (MP) method.

The aim of the method in Ref. 4 was to detect rectangular road signs such as motorway road signs. The detection was based on finding rectangular shapes using the Hough Transform in the edge map extracted from the grey level image of the road. Therefore, this method strongly depended on the quality of the edge detection stage.

Researchers have put efforts in finding road signs, however, simply finding and capturing images of the road signs and storing these for later viewing is not a good solution since you either have to record all the road signs that you found in the process or to spend much longer time to analyze what the road sign was. The information on the sign may not be of good quality when the sign was first detected and no previous research discussed the quality of road sign images. The poor quality of the road sign image may be due to the road sign is too far to read, too close to have the whole sign, or too bright/dark to distinguish the characters on the sign.

The aim of this project was to detect and track road signs in the live image sequence and pick up the best moment to record the image of the road sign so that automated Optical Character Recognition (OCR) procedure could be carried easily later to determine the information on a road sign. The information could then be stored instead of images. In this way, a symbolic representation of a road sign can be formed and stored for later viewing by the driver.

We described in this paper a procedure based on a single camera vision system, which can detect and extract road signs from a sequence of images. The following sections described the main aspects of the procedure, namely road sign detection, tracking and quality computation.

Road Sign Extraction

The usage of colour (blue, green, yellow, and red) for the road sign detection corresponds to the way in which they are usually picked up by human eyes. In that case, we can classify, as in [2], each pixel as 1 (part of road sign) or 0 according to its hue, saturation and intensity. Hence, a road sign corresponds to a high density of pixels that have the "same colour". Afterwards, we can extract each region formed by this method using a growing method. Then, the accuracy of the position of the road sign is improved using morphological opening. Finally, if the region is big enough it could be a good candidate to be a road sign.

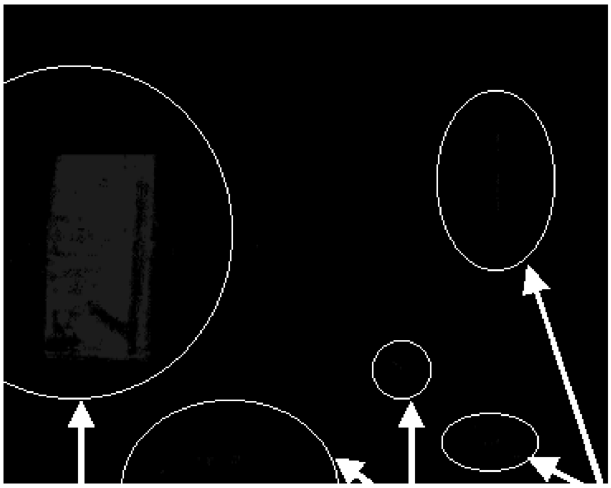
Figure 1. Typical road image, with the detected road sign.

The use of colour[1,2,3] provides the fastest way to find a coloured sign in a road scene. The case of a white road sign will be discussed later. This method reduces considerably the search region to a small part of the image which can be validated to be a road sign in a further stage. In fact, the use of colour for image segmentation corresponds to the way road signs are usually picked up by drivers.

In order to perform this detection, a colour model has to be chosen. Considering the way in which human beings perceive colour, the HSI model is chosen.

The colour for the road signs can be measured and calibrated against the different illumination conditions. Motorway road signs in UK, for example, can now be consideredd as "deep blue" according to its hue, saturation and intensity. The hue must be between 200° and 230° in an angular scale, this range corresponds to an acceptable blue for a road sign. Then the saturation must lie between 50% and 90% to differentiate the blue of the road sign from the blue of the sky. Finally the intensity, which gives an idea of the brightness of the pixel should lie between 25% and 90% to avoid picking up completely black or white pixels.

Hence each pixel within the region of interest could be classified considering that they are more likely to be part of a road sign. We ended up with a "colour map" shown in figure 2, which represented the result of the application of this method and where we could identify two types of information: high density regions, which were more likely to be road signs and isolated peaks, which had to be considered as noise.



Road sign

Isolated small regions

Figure 2. Colour grouping result.

The same technique could be used to detect road signs of other background colours.

The case of a white road sign was slightly different, since white was not considered as a colour: there was no range of hue corresponding to white. In fact, a white could be considered if the intensity was higher than 90% and the saturation was lower than 10%, for any value of hue. Therefore, these values can be used in order to pick up white road signs. However, the numbers of white objects in a road scene were more common than deep green, deep blue or deep red objects. For instance, if the sky is white, it will be picked up by the system. To reduce the number of wrong detections, the region of interest was found using some segmentation techniques to avoid processing the sky.

A simple way to know if the region found was more likely to be a road sign was to check if something is written on it. One way to know if letters could be found inside a region is to cross correlate the region with letters taken from a database.[4] The use of vowels seemed to be a good idea, but since the car was moving, the road sign (indeed the letters) appeared in different sizes, and rotation could appear as well.

Several methods could be used such as OCR algorithms based on neural networks, but this was not the purpose of this research. In fact, a realistic real-time vision system was not able to use an OCR algorithm on each frame grabbed from the camera. Therefore, fastest methods should be used to know if some text was present inside this region.

The method used in this paper was based on the colour map defined earlier. In fact, the holes left from the colour-growing region were usually characters or symbols printed on the sign, which were of the same color, either in bright (on blue and green road signs) or dark (on yellow and white road signs).

Considering the first class of road signs, some ratios can be set from the picture:

- α: number of "bright" pixels within the holes over the total number of pixels in that region,
- β: number of "bright" pixels within the holes over the number of pixels within the holes.

If the first ratio is greater than 10% and the second one greater than 30%, then the region is a good candidate to be considered as a road sign.

Road Sign Tracking

A tracking algorithm was used in to improve the robustness of the previous detection procedure. Once a possible road sign has been located, it was possible to estimate its position in later frames, knowing the speed and the direction of the vehicle. If nothing is found in the predicted area then there was probably a wrong detection. For instance, the back of a car or a truck would most likely be moving with approximately the same speed as the camera whereas a stationary road sign would appear closer and closer in a sequence of images. The tracking procedure was also used to reduce the computational time of the overall procedure.[2]

The main purpose for us to a tracking procedure was to link road signs from one frame to another in run time.

The tracking procedure applied in this research has two modes, namely *coarse mode* and *fine mode*. In the coarse mode, detection algorithm was only applied within key frames (for instance every 5 frames). Within the inner frames, the detection algorithm was only applied in the regions that have been predicted with the tracking process. The advantage of this solution was that the system did not have to process the whole field of every frame, therefore the computational time is considerably reduced. This was an important factor for a real time vision system. However, in this mode the system could fail if the prediction is not reliable. In fact, most of the wrong prediction is due to a sudden change in the velocity of the car.

The implementation of the fine mode is much more straightforward but also more expansive in terms of computation. In this mode, every frame was monitored for road sign detection within the region of interest. From each road sign within the current frame, a prediction of the position in the next frame was carried out and compared with the closest object found in the next frame. If the two regions matched, we can consider that the region found was the same road sign. This stage was also called "linking process" because it allows us to follow a road sign in a sequence of images and to observe its motion as well. Finally if the object tracked "does not move like a road sign" (its motion does not correspond to the expectations) then there was probably a false detection. Since all the frame is processed, the risk of losing a sign is considerably reduced. Because the motion of the road sign was observed in this mode, we could switch back to the coarse mode if the motion became stable again.

The feature points used for the tracking procedure were the corners of the detected road sign region. The vehicle is assumed to be moving with approximately constant speed v on an approximately straight road with constant slope. Any difference to this hypothesis was taken into account as an error source. The road sign is assumed to be orthogonal to the road direction, and therefore parallel to the image plane.

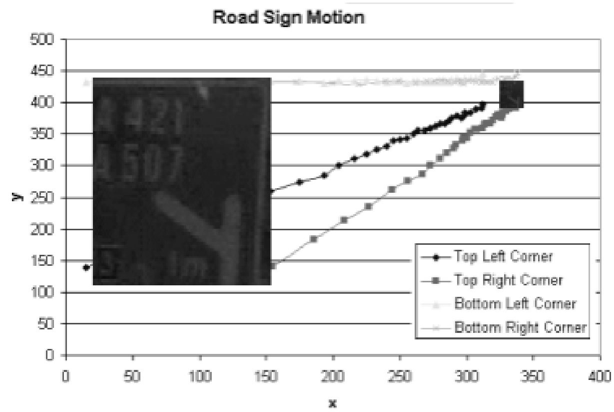


Figure 3.Tracking result

Choice of the Best Road Sign Images

If the road sign was too far or too close to the vehicle, the driver or an OCR algorithm would not be able to read it even if it has been successfully detected. In fact, what we really need to know was, in a sequence of images what image was supposed to be the best candidate to be stored on the hard drive for later display to the driver or for OCR purpose. A *probability* was given to the road sign and updated according to its "behaviour" in the sequence. For instance, its probability is increased if the road sign is following the rules stated in the tracking algorithm. Hence, the road sign will simply be stored when its probability reaches a maximum.

Once one or more signs have been found and followed, the best time to store them on the hard drive has to be found. In daylight conditions this will usually mean they are as close to the vehicle as possible without any of them passing out of the camera's field of view. At night, a different distance might be largely due to the effect of the headlight. However, all the results reported in this paper has been done in daylight.

We formed a combined parameter, namely *probability*, to assign to the road signs that we found in every image as soon as they were picked up. The probability was then updated with respect to different features and gives us an idea of the legibility of the road sign. The use of this parameter (probability) for road sign detection allows us to store on the hard drive the "best" road sign.

This parameter could be formed from different features: the size, the motion, and the text. Some other features can also be included such as in Ref. 5, with the use of wavelets and neural network to know if the road sign has some text on it and how legible it is. Another feature, which can be included, is the presence of straight lines (borders) in the road sign as in Ref. 4. The use of a shape analyser can be part of the probability since road signs, which have been detected were rectangles. A weight coefficient can also be assigned to each feature enabling them to be prioritised. For instance, if we consider that the presence of text is more important or more reliable than straight lines we can put more weight in the first feature.

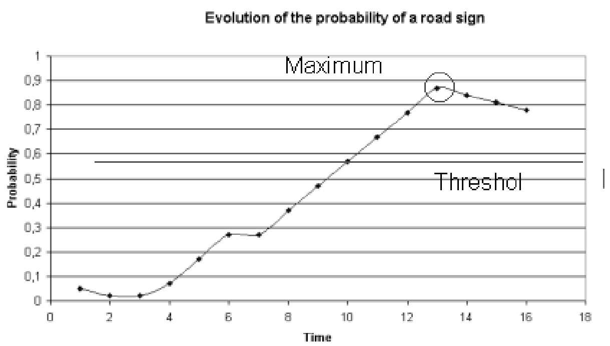


Figure 4. Evolution of the probability of a road sign.

In this research, we used the width and height of the region of detected sign, and texture content as the components of the probability for our purpose. The width and height were calculated according to the corner positions obtained in the previous stages whilst the texture content information came from the size of the hole area within the region of the road sign. The algorithm make the probability to increase when the width and height increased linearly, also when the ratio of the text content (the ratio of the size of the hole region to the size of the road sign region) increased. The ration of the text content should become constant and the size of the hole region should increase at the rate that approximately equal to the width increasing rate times height increasing rate.

Only the road sign, which has the highest probability is stored on the hard drive. This road sign is supposed to be the best candidate for the application of an OCR algorithm.

Experimental Results and Discussion

The following figures show the results of the road sign detection algorithm on static images. The *search regions* are highlighted in green and the *found regions* are highlighted in red.

In the figure 5(a), two of the three road signs were detected but not confirmed (no red rectangle). They are too small to be considered. In figure 5(b) (a closer image) the green and the blue road signs are picked up correctly by the system. The white road signs are not considered.

In the second example, shown in figure 6, the blue car was mistaken as the candidate of a road sign but was later discarded.

In the first figure, the found region does not cover all the road sign. However, in a later image, the whole road sign has been correctly detected. In this example, we can see that the segmentation is well performed between the green of the trees and the green of the road sign.

The road sign has been correctly detected, even with a complicated combination of the road signs

(a)

(b)

Figure 5.

(a)

(b)

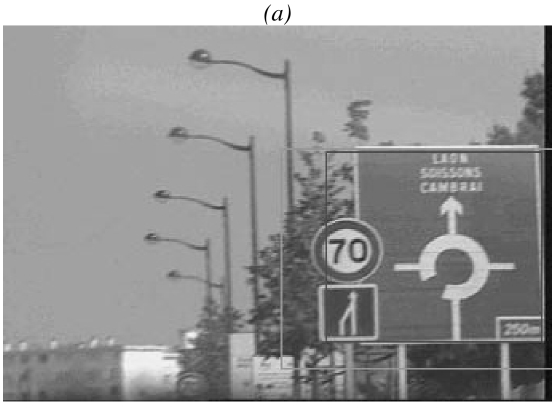
Figure 6

(a)

(b)

Figure. 7

(a)



(b)

Figure 8.

Conclusion

The results of applying proposed technique to the real image sequences taken from motorway in France and UK show that road signs are successfully detected and tracked. The recorded images of road signs are of the best quality in the sequence. Computational speed is fast and the algorithm should be able to be adapted to real-time implementation.

References

1. W. Ritter, F. Stein, and R. Janssen, Traffic Sign Recognition Using Colour Information, Mathl. Comput. Modelling Vol. 22, No. 4-7, 149-161, (1995).
2. G. Picciola, E. De Micheli, P. Parodi, and M. Campani, Robust method for road sign detection and recognition, Image and Vision Computing Journal Vol. 14, 209-223, (1996).
3. S. H. Hsu and C. L. Huang, Road sign detection and recognition using matching pursuit method, Image and Vision Computing Journal Vol. 19, 119-129, (2001).
4. S. Challands and S. Dickson, Timely capture of road signs via OCR, Cranfield University, MSc Thesis, (1999).
5. H. Li, D. Doermann, and O. Kia, Automatic Text Detection and Tracking in Digital Video, Language and Media Processing Laboratory, Institute for Advanced Computer Studies, University of Maryland, Mathematical and Computational Sciences Division, National Institute of Standards and Technology, (1998).

Biography

Eric Gambardella received his B.S. degree in Computer Science from the Universitede Technologie de Compiegne (UTC) in 2000 and a MSc. in Software Engineering for Digital Signal Processing from Cranfield University. He is now working for AXYS-Consultants.

Shan Fu received his B.S. degree in Electronic Engineering from the North-Western Polytechnic University in China and a Ph.D. from Heriot-Watt University in UK. He is now a lecturer in DSIP at Cranfield University. His research interest is in Computer Vision and its applications in Engineering.