

TRAFFIC LIGHTS DETECTION IN ADVERSE CONDITIONS USING COLOR, SYMMETRY AND SPATIOTEMPORAL INFORMATION

George Siogkas, Evangelos Skodras and Evangelos Dermatas

Electrical and Computer Engineering Department, University of Patras, Rio, Patras, Greece

Keywords: Traffic Light Detection, Computer Vision, Fast Radial Symmetry, Adverse Conditions, Spatiotemporal Persistency.

Abstract: This paper proposes the use of a monocular video camera for traffic lights detection, in a variety of conditions, including adverse weather and illumination. The system incorporates a color pre-processing module to enhance the discrimination of red and green regions in the image and handle the “blooming effect” that is often observed in such scenes. The fast radial symmetry transform is utilized for the detection of traffic light candidates and finally false positive results are minimized using spatiotemporal persistency verification. The system is qualitatively assessed in various conditions, including driving in the rain, at night and in city roads with dense traffic, as well as their synergy. It is also quantitatively assessed on a publicly available manually annotated database, scoring high detection rates.

1 INTRODUCTION

Advanced Driver Assistance Systems (ADAS) are new emerging technologies, primarily developed as automated advisory systems to enhance driving safety. One of the most challenging problems for such systems is driving in urban environments, where the visual information flow is very dense and can cause fatigue, or distract the driver. Aside from moving obstacles like cars and pedestrians, road signs and traffic lights (TL) significantly influences the reactions and decisions of the driver, which have to be made in real time. The inability to make correct decisions can lead to serious, even fatal, accidents. Traffic light violations in particular, are one of the most common causes of road crashes worldwide.

This is where ADAS can provide help that may prove life-saving. A very big portion of such systems is based on visual information processing using computer vision methods. Whether in the form of completely automated driving systems, using only visual information, or in the form of driver alert systems, technology can provide crucial assistance in the efforts to reduce car accidents. Both the aforementioned systems have to include a reliable and precise Traffic Light Detection (TLD) module, so that accidents in intersections are mitigated.

The idea of using vision for ADAS in urban environments so that TLD can be achieved was first introduced in the late 90's, by (Loy and Zelinsky, 2002) who proposed a computer vision based Stop & Go algorithm using a color on-board camera. However, the use of computer vision for ADAS bloomed in the next decade, as computer processor speeds reached a point that enabled real-time implementation of complex algorithms. The work of (Lindner et al., 2004) proposes the fusion of color cameras, GPS and vehicle data to increase the robustness of their TLD algorithm, which is followed by a tracking and a classification module. The TLD part uses RGB color values, texture and shape features. In the HSV space, color thresholding followed by a Gaussian filtering process and a verification of TL candidates is the approach of (In-Hak et al., 2006) to detect TLs in crossroads. A similar approach is followed by (Yehu Shen et al., 2009), based on HSV images and a Gaussian distribution-based model acquired by training images. A post processing phase utilizes shape and temporal consistency information to enhance the results. A more straight-forward process is proposed by (M. Omachi and Omachi, 2009), who locate candidate regions in the normalized RGB color space and validate the results using edge and

symmetry detection. Color information has been ignored by (de Charette and Nashashibi, 2009a), who use grayscale spot detection followed by an Adaptive Template Matcher, achieving high precision and recall rates in real time. Their results were tested thoroughly and compared to the results of an AdaBoost method using manually annotated videos in (de Charette and Nashashibi, 2009b). The problem of TLD in day and night scenes has been addressed by (Chunhe Yu et al., 2010) using RGB thresholding and shape dimensions to detect and classify lights in both conditions.

By reviewing related literature to date, the following conclusions can be drawn about vision based TL detection:

- Researchers have not always used color information; when they do, they prefer either the HSV, or RGB color spaces.
- Many groups have used empirical thresholds that are not optimal for all possible driving conditions (i.e. shadows, rain, and night). Generally, adverse conditions are not very frequently addressed.
- Symmetry is frequently used, either with novel detection techniques, or with the well-known Hough transform, but never with the fast radial symmetry transform (Loy and Zelinsky, 2003).
- Traffic lights candidate detection is followed by a validation process, required to exclude false positive results. The use of TL models, tracking, or both is the most common solution.
- Apart from (Robotics Centre of Mines ParisTech, 2010) that provides a publicly available annotated database of on-board video frames taken in Paris, to the best of our knowledge, there are no other annotated databases for traffic lights detection.

This paper presents a TLD algorithm inspired by the approaches followed for road sign detection, by (Barnes and Zelinsky, 2004), (Siogkas and Dermatas, 2006) and (Barnes et al., 2008). The Fast Radial Symmetry (FRS) detector of (Loy and Zelinsky, 2003) is employed in the referenced approaches, to take advantage of the symmetrical geometry of road signs. The symmetry and color properties are similar in road signs and traffic lights, so these approaches can be a good starting point. The goal of the system is to provide a timely and accurate detection of red and green traffic lights, which will be robust even under adverse illumination or weather conditions.

The proposed system is based on the CIE- $L^*a^*b^*$ color space (Illuminant, 1978) exploiting

the fact that the perceptually uniform a^* coefficient is a color opponency channel between red (positive values) and green (negative values). Therefore, it is suitable for distinction between the two prominent classes of TLs. An image processing phase comprising 4-connected neighborhood image flood-filling on the positive values of a^* channel is then applied, to ensure that red traffic lights will appear as filled circles and not as black circles with a light background. The fast radial transform is then utilized to detect symmetrical areas in a^* . The proposed system has been tested in various conditions and has been qualitatively and quantitatively assessed, producing very promising results.

The rest of the paper is organised as follows: Section 2 presents an overview of the proposed system, describing in depth every module and its functionalities. Section 3 presents experimental results in both normal and adverse conditions, and finally, section 4 discusses the conclusions drawn from the experiments and suggests future work.

2 PROPOSED SYSTEM

The hardware setup for the proposed system is similar to most of the related applications. The core of the system is a monocular camera mounted on an elevated position on the windshield of the moving vehicle. The video frames shot by the camera are processed by three cascade modules. The first one is the pre-processing module, with a goal to produce images in which red TLs will appear as very bright circular blobs and green TLs will appear as very dark circular blobs. The image obtained by the pre-processing module is used in the traffic light detector, which comprises a FRS transform for various radii, followed by a detection of multiple local maxima and minima in the top part of the frame. The last module applies a spatiotemporal persistency verification step to keep those candidates that appear in multiple frames, thus minimizing false positives.

Summing up, the proposed algorithm consists of the following steps:

- 1) Frame acquisition.
- 2) Image pre-processing:
 - a) Convert RGB to $L^*a^*b^*$.
 - b) Enhance red and green color difference.
 - c) Fill holes in enhanced image
- 3) TL candidate detection:
 - a) Radial symmetry detection.
 - b) Maxima/minima localization.

- 4) TL candidate verification:
 a) Spatiotemporal persistency check.

The steps described above are shown in Figure 1.

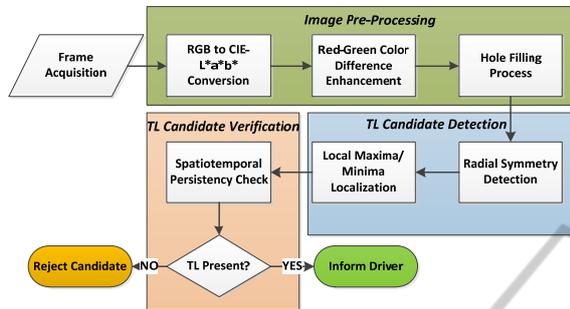


Figure 1: Proposed TL detection algorithm.

2.1 Pre-processing Module

Before processing each frame to detect TL candidates, the contrast between red and green lights and their circularity are enhanced. Ideally, red TLs should be very bright circular blobs and green TLs should be very dark circular blobs. In this direction, the perceptually uniform a^* channel of the CIE- $L^*a^*b^*$ color space, which assigns large positive and negative values to red and green pixels respectively, is utilized. To further enhance the discrimination between red and green, the pixel values of the a^* channel are multiplied by the pixel luminosity (L channel), to produce a new channel, called RG hereafter, defined as

$$RG(x, y) = L(x, y) \times a(x, y), \quad (1)$$

where x, y are the pixel coordinates.

This transformation results in a further increase of the absolute value of the pixels belonging to TLs, as they also tend to have large luminosity values. On the other hand, green objects with very large absolute values of a^* like tree leaves, do not appear so bright, so they are not affected on the same degree. The same applies for most red objects, like red roofs. The aforementioned process is demonstrated in Figure 2. While the green TLs are still dark when multiplying L by a^* (Figure 2d), the leaves are not as dark as in the a^* channel (Figure 2c).

Ideally, the image transformation from the RGB to the CIE- $L^*a^*b^*$ color space would be enough to achieve the aforementioned goal. However, real world video sequences containing TLs often produce a “blooming effect”, especially in the case of red TLs, as shown in Figure 3a. The cause of this blooming effect can be twofold: i) red lights often include some orange in their hue, while green TLs

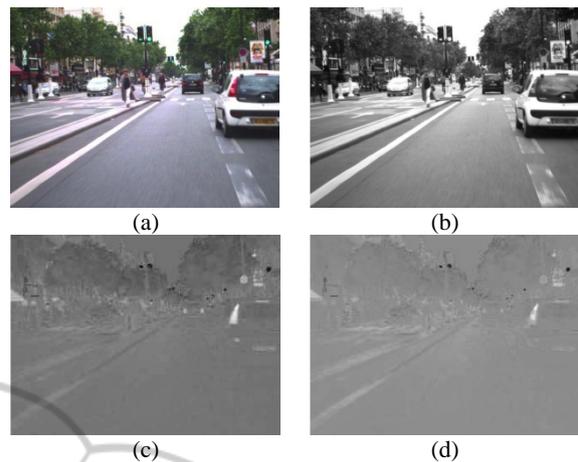


Figure 2: Red-green discrimination enhancement. (a) RGB image, (b) L channel, (c) a^* channel, (d) L, a^* multiplication result (RG channel).

include some blue. ii) The dynamic range of cameras may be sensitive to very bright lights; consequently saturated regions appear in their inner area.

To tackle this problem, we propose to combine the red and green areas with the yellow and blue ones, respectively. This is accomplished by calculating the product of L and b^* (YB channel):

$$YB(x, y) = L(x, y) \times b(x, y) \quad (2)$$

where x, y are the pixel coordinates and then adding the result to the RG channel to produce the RGYB image:

$$\begin{aligned} RGYB(x, y) &= RG(x, y) + YB(x, y) \\ &= L(x, y) \times (a(x, y) + b(x, y)). \end{aligned} \quad (3)$$

The results of this process are demonstrated in Figure 3. The “blooming effect” is shown in the red TL on the right of the picture in Figure 3b. While it would be expected that the value of a^* to be positive, it appears to be negative (green). The redness is observed only in the perimeter of the TL and could lead the FRS algorithm to detect two symmetrical shapes with the same centre (a large bright one and a smaller dark one). However, by adding the YB channel (Figure 3c) to the RG result, this problem is handled, as shown in Figure 3d.

The blooming effect, however, still remains a problem in the case of night driving, so an additional step is necessary. This step is a grayscale 4-connected neighbourhood image filling (Soille, 1999) of the holes in both the bright and the dark areas of the image. More specifically, the RGYB image is thresholded to produce two images: one with red and yellow pixels ($RGYB > 0$) and one

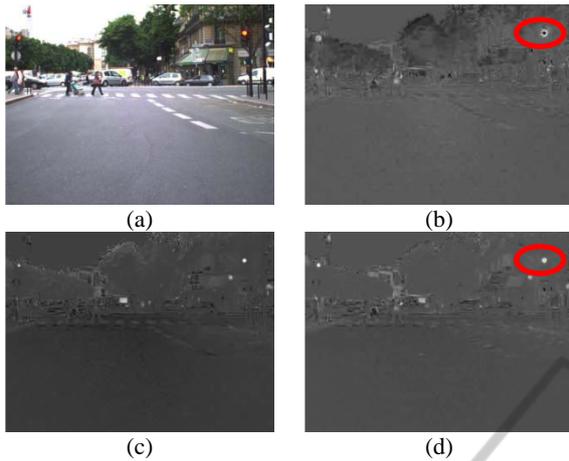


Figure 3: Handling the “blooming effect” in the morning. (a) RGB image, (b) RG channel, (c) YB channel, (d) RGYB channel.

with green and blue pixels ($RGYB < 0$). The holes in both images are then filled, using the method mentioned above, and then they are added to produce the final result. This process and its results (denoted by the red ellipses) are demonstrated in Figure 4. The green TLs in the scene are shown as big black circles with a concentric smaller, lighter circle inside them (Figure 4b). This effect has been eliminated after the aforementioned filling process, as shown in Figure 4c.

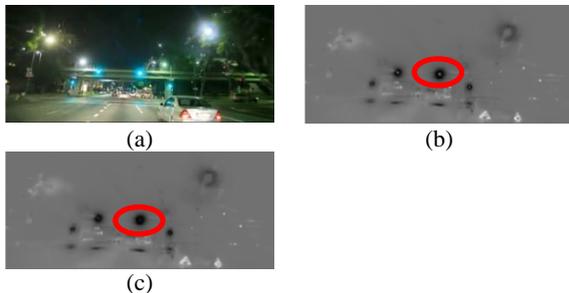


Figure 4: Handling the “blooming effect” at night. (a) RGB image, (b) RGYB channel, (c) RGYB channel after filling process.

2.2 Radial Symmetry Detection

The algorithm for FRS detection was first introduced by (Loy and Zelinsky, 2002) and improved in (Loy and Zelinsky, 2003). Its main feature is detecting symmetrical blobs in an image, producing local maxima in the centres of bright blobs and local minima in the centres of dark blobs. The only parameters that have to be defined for this process are a radial strictness factor, a , and the radii that will be detected.

As mentioned in section 2.1, the image transformed by the FRS algorithm includes red and green color opponency. This property makes it appropriate for the FRS transform. Some examples of the implementation of the FRS transform (for radii from 2 to 10 pixels with a step of 2 and $a=3$) to pre-processed frames of various videos are reported in Figure 5.



Figure 5: Fast radial symmetry transforms (right column), at day and night time (left column). Dark spots denote green TLs, light spots denote red TLs. Local maxima/minima detection is performed above the yellow line (which depends on the camera placement).

The results of the FRS transform that are demonstrated in Figure 5 show that the centres of TLs are always within the most voted pixels of the frame, as long as their radii are included in the search. This means that once the FRS algorithm has been applied, the following process is a non-maxima suppression to detect red TLs and a non-minima suppression to detect green TLs. However, since the nature of the TL detection problem allows it, only the conspicuous region of the image, i.e. the upper part, will be used for the suppression. Since the on-board camera was not placed identically in all the videos used, the selection of the part of the image that will be used for the suppression must be defined during the system calibration. This could be achieved automatically using horizon detection techniques, but this is not necessary because the areas where a TL could appear can be predicted during the camera setup. In the examples of Figure 5, the suppression took place above the yellow line.

Up to 5 local maxima and 5 local minima are selected, given that they lie within the range of numbers greater than half the value of the global maximum. Once the TL candidates are produced, the

TL areas are denoted by a rectangle with coordinates that are determined by the detected centre coordinates, the radius and the color of the TL. More specifically, the annotation rectangle for a green TL starts at 6 radii up and 1.5 radii to the left of the TL centre and has a width of 3 radii and a height of 7.5 radii. Similarly, the annotation rectangle for a red TL starts at 1.5 radii up and 1.5 radii to the left of the TL centre and has the same height and width as above.

2.3 Candidate Verification

The last module of the proposed system is a TL candidate verification process. This step is vital for the minimization of false positives that could be caused by numerous artefacts that resemble a TL. Road signs, advertisements etc. can be misinterpreted as TLs, because their shape and color properties might look alike. However, such artefacts usually don't appear radially symmetrical for more than a few frames, while the symmetry and color of TLs are more robust to temporal scale variations. Hence, most of them can be easily removed if the condition of multiple appearance of a TL in successive frames is met.

In order to satisfy spatiotemporal persistency, the proposed method assumes that a TL will appear in the top voted candidates for symmetrical regions in a sequence of frames (temporal persistency), so its centre is expected to leave a track of pixels not far from each other (spatial persistency). Such a result is shown in Figure 6, where the trails left by the centres of two green TLs over a period of 30 frames are denoted by green dots. The sparse dots on the left belong to the other detected symmetrical objects and do not fulfil the persistency criterion.



Figure 6: Trails left by the centres of two green TLs.

A persistent high FRS value in a small area for at least 3 out of 4 consecutive frames is the post-processing criterion for characterizing a pixel as being the centre of a TL. An example of the aforementioned process is demonstrated in Figure 7, where the first column shows frames 1, and 2 and the second column contains frames 3, and 4. The red rectangles denote a red TL candidate and the yellow ones denote a green TL candidate.

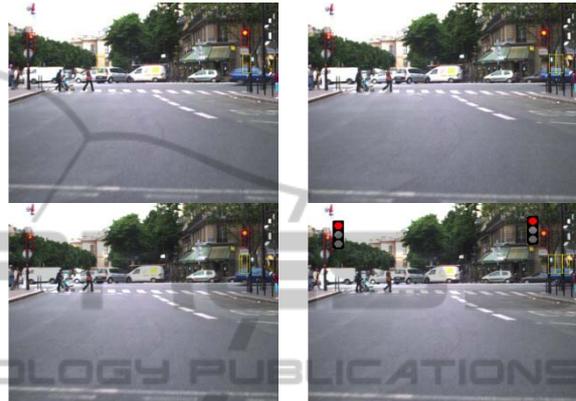


Figure 7: Four consecutive frames, TL candidates annotated by rectangles. Non persistent candidates are dismissed. Candidates persistent for 3 out of 4 consecutive frames get verified.

3 EXPERIMENTAL RESULTS

The assessment of the effectiveness of the proposed system is twofold.

First, a quantitative assessment is made, using the publicly available, manually annotated video sequence of (Robotics Centre of Mines ParisTech, 2010). The video sequence comprises 11179 frames (8min 49sec) of 8-bit RGB color video, with a size of 640x480 pixels. The video is filmed in a C3 vehicle, travelling at speeds less than 50km/h, and the camera sensor is a Marling F-046C, with a frame rate of 25fps and a 12mm lens, mounted behind the rear-view mirror of the car.

Secondly, extensive qualitative testing is carried out, using video sequences of different resolutions shot with different cameras, under adverse conditions in day and night time, in various countries, downloaded off the internet (mainly from the YouTube site), or 720x576 pixels videos at 25fps filmed in Greek roads, using a Panasonic NV-GS180 video camera. The purpose of this experimental procedure is to examine the system's robustness in changing not only the driving, weather and illumination conditions, but also the quality of the

videos and the camera setup.

The proposed system was implemented in Matlab and was tested on a computer with a Core 2 Quad CPU at 2.83GHz, and 4GB of memory. The code did not use any parallelization. The processing times achieved were directly affected by the resolution of the videos and they fluctuated from 0.1ms to 0.5ms per frame.

As far as the parameters used for the proposed algorithm are concerned, the radii for the FRS in the experiments were 2,4,6,8, and 10 pixels. The radial strictness was 3, and for the spatiotemporal persistency to hold, a candidate must appear in at least three out of four consecutive frames, in an area of a radius of 20 pixels.

3.1 Quantitative Results

The quantitative results of the TL detector are estimated following the instructions given in (Robotics Centre of Mines ParisTech, 2010), for their publicly available, manually annotated video sequence. These instructions stated that from the 11179 frames included in the video sequence, only 9168 contain at least one TL. However, not all the TLs appearing in these frames are used: yellow lights (58 instances) are excluded, as well as many lights which are ambiguous due to heavy motion (449 instances). Apart from these, 745 TLs that are not entirely visible, i.e. are partially outside the frame, are also excluded. Eliminating all these TL instances, a total of 7916 visible and unambiguous red or green TLs remain. The total distinct TLs that comprise these 7916 instances are 32.

The results scored by the proposed system are shown calculated following the same rule as (de Charette and Nashashibi, 2009a), which states that if a TL is detected and recognized once in the series of frames where it appears, then it is classified as a true positive. Hence, a false negative result is a TL never detected in its timeline. Using these definitions, the proposed algorithm scores a detection rate (recall) of 93.75%, detecting 30 of the total 32 TLs. The 2 missed traffic lights were green and appeared for 49 and 167 frames in the scene. In both cases, our system detected the other green lights appearing in the same scene, so the goal of informing the driver can be achieved. Some detection examples from the aforementioned database are given in Figure 8. The total number of false positive detections is 1481, of which 1167 are false red TL positives and 314 false green TL detections. The red false positive instances concern 12 different objects mistaken as red TLs, while the number of different objects mistaken as

green TLs is 7. This means that the precision of our system is 61.22%. This false positive rate could be significantly reduced by using morphological cues, like the templates utilized by (de Charette and Nashashibi, 2009a) and (de Charette and Nashashibi, 2009b), who report precision and recall rates of up to 98% and 97% using temporal matching. A more direct comparison between our system and the two aforementioned approaches is not feasible though, due to the lack of details in the evaluation process.



Figure 8: TL detection results in daytime driving. A false positive red TL example is shown in the bottom left image and a false negative is shown in the bottom right.

3.2 Qualitative Results

The experimental results reported in section 3.1 show that the system appears effective and robust in the case of urban driving in good weather and illumination conditions. The problem of false positives is not so persistent and could be further improved if, as already mentioned, a TL model is used for a final verification, or if a color constancy module is introduced. However, the great challenge of such systems resides in more demanding driving conditions, including driving under rainy conditions and driving at night time.

For this reason the proposed system is also tested in such conditions, so that its robustness and resilience can be examined. Various video sequences shot from on-board cameras all around the world have been gathered from the Internet. The ultimate goal is to construct a database of challenging conditions driving videos, which will be annotated and freely distributed in the future.

3.2.1 Driving Under Rainy Conditions

The first challenging case in ADAS is driving in rainy conditions. The difficulty present is that

raindrops on the windshield often distort the driving scene and could cause the various objects to appear disfigured. Another problem in rainy conditions is the partial obstruction by the windshield wiper in various frames. For these reasons, every vision based system should be tested under rainy conditions, as the results produced may vary a lot from the ones achieved in normal weather. Some examples of successful TL detections in rainy conditions are shown in Figure 9.



Figure 9: Successful TL detection results in rainy conditions.

The presence of a TL for the predefined number of frames is annotated by the inclusion of a manually drawn TL template next to the detected TL. Some false positive examples are annotated, but have not appeared in previous frames, i.e. the green rectangle in the last image on the left column of Figure 9. The false negatives in the same picture (green TL on the top) and in the first picture of the right column of Figure 9 (green TL on the right) have not reached 5 consecutive detections yet and are recognised in a later frame.

3.2.2 Driving at Night

The second important category of adverse driving conditions is night driving. The difficulty of the situation relies largely on the environment and the mean luminosity of the scene. If the environment does not include excessive noise like for example dense advertisements or other lighting sources, the proposed system performs well, even in urban driving situations. Successful detections in night

driving scenarios are presented in Figure 10. Most TLs are successfully detected, even when their glow makes it very difficult to distinguish shape and morphology cues.



Figure 10: TL detection results in urban night driving.

3.2.3 Known Limitations of the TLD

A very common question when dealing with computer vision applications is whether there are problems that they cannot solve. The proposed method is by no means flawless and can produce persistent errors in some situations. The main problems that can be pinpointed are the following:

i) The system produces some false positive results that cannot be easily excluded, unless it is used in correlation to other computer vision modules like a vehicle detector or a road detector. An example of such false positives is illuminated vehicle tail lights, or turn lights, as in Figure 11a. This image also includes a false positive result that is excluded in next frames. ii) The proposed system fails completely in cities like New York (Figure 11b), where the visual information is extremely dense and the TLs are lost in the background.



Figure 11: Examples of temporary or permanent failures of the proposed system.

4 CONCLUSIONS

In this paper, we have proposed a novel automatic algorithm for traffic lights detection using a monocular on-board camera. The algorithm uses

color, symmetry and spatiotemporal information to detect red and green traffic lights in a fashion resilient to weather, illumination, camera setup and time of day. It utilizes a CIE-L*a*b* based color space with a holes filling process to enhance the separability of red and green traffic lights. A fast radial symmetry transform is then used to detect the most symmetrical red and green regions of the upper part of the image, producing the TL candidates. Finally, a spatiotemporal persistency criterion is applied, to exclude many false positive results. The algorithm has been experimentally assessed in many different scenarios and conditions, producing very high detection rates, even in very adverse conditions.

Future work will be directed towards embedding a tracking module to the algorithm to minimize the false negative results and a color consistency module to further reduce false positives. Furthermore, the combination of the TL detector with other ADAS modules like vehicle, sign and road detection will be explored, so that a complete solution for driver assistance is proposed.

REFERENCES

- Barnes, N., & Zelinsky, A. (2004). Real-time radial symmetry for speed sign detection (pp. 566–571). Presented at the 2004 IEEE Intelligent Vehicles Symposium, IEEE. doi: 10.1109/IVS.2004.1336446
- Barnes, N., Zelinsky, A., & Fletcher, L. S. (2008). Real-time speed sign detection using the radial symmetry detector. *Intelligent Transportation Systems, IEEE Transactions on*, 9(2), 322–332.
- Chunhe Yu, Chuan Huang, & Yao Lang. (2010). Traffic light detection during day and night conditions by a camera (pp. 821–824). Presented at the 2010 IEEE 10th International Conference on Signal Processing (ICSP), IEEE. doi:10.1109/ICOSP.2010.5655934
- de Charette, R., & Nashashibi, F. (2009a). Real time visual traffic lights recognition based on Spot Light Detection and adaptive traffic lights templates (pp. 358–363). Presented at the 2009 IEEE Intelligent Vehicles Symposium, IEEE. doi:10.1109/IVS.2009.5164304
- de Charette, R., & Nashashibi, F. (2009b). Traffic light recognition using image processing compared to learning processes (pp. 333–338). Presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009. IROS 2009, IEEE. doi:10.1109/IROS.2009.5353941
- Illumination, I. C. on. (1978). *Recommendations on uniform color spaces, color-difference equations, psychometric color terms*. Bureau central de la CIE.
- In-Hak, J., Seong-Ik, C., & Tae-Hyun, H. (2006). Detection of Traffic Lights for Vision-Based Car. *Traffic*, 4319/2006, 682–691. doi:10.1007/11949534_68
- Lindner, F., Kressel, U., & Kaelberer, S. (2004). Robust recognition of traffic signals (pp. 49–53). Presented at the 2004 IEEE Intelligent Vehicles Symposium, IEEE. doi:10.1109/IVS.2004.1336354
- Loy, G., & Zelinsky, A. (2002). A fast radial symmetry transform for detecting points of interest. *Computer Vision—ECCV 2002*, 358–368.
- Loy, G., & Zelinsky, A. (2003). Fast radial symmetry for detecting points of interest. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 959–973.
- Omachi, M., & Omachi, S. (2009). Traffic light detection with color and edge information (pp. 284–287). Presented at the 2nd IEEE International Conference on Computer Science and Information Technology, 2009. ICCSIT 2009, IEEE. doi:10.1109/ICCSIT.2009.5234518
- Robotics Centre of Mines ParisTech. (2010, May 1). Traffic Lights Recognition (TLR) public benchmarks [La Route Automatisée]. Retrieved October 6, 2011, from <http://www.lara.prd.fr/benchmarks/trafficlightrcognition>
- Siogkas, G. K., & Dermatas, E. S. (2006). Detection, Tracking and Classification of Road Signs in Adverse Conditions (pp. 537–540). Presented at the IEEE Mediterranean Electrotechnical Conference, MELECON 2006, Malaga, Spain. doi:10.1109/MELCON.2006.1653157
- Soille, P. (1999). *Morphological image analysis: principles and applications with 12 tables*. Berlin [u.a.]: Springer.
- Yehu Shen, Ozguner, U., Redmill, K., & Jilin Liu. (2009). A robust video based traffic light detection algorithm for intelligent vehicles (pp. 521–526). Presented at the 2009 IEEE Intelligent Vehicles Symposium, IEEE. doi:10.1109/IVS.2009.5164332