



# An application of a bio-inspired visual attention model for the localization of vehicle parts



Ana-Maria Cretu<sup>a,b,\*</sup>, Pierre Payeur<sup>b</sup>, Robert Laganière<sup>b</sup>

<sup>a</sup> Université du Québec en Outaouais, 101 Saint-Jean-Bosco, Gatineau, QC J8X 3X7, Canada

<sup>b</sup> University of Ottawa, 800 King Edward, Ottawa, ON K1N 6N5, Canada

## ARTICLE INFO

### Article history:

Received 5 November 2013

Received in revised form 27 January 2015

Accepted 20 February 2015

Available online 16 March 2015

### Keywords:

Bio-inspired systems

Visual attention

Automotive vehicles

Vehicle part identification

Image classification

Active shape models

## ABSTRACT

The automated servicing of vehicles is becoming more and more a reality in today's world. While certain operations, such as car washing, require only a rough model of the surface of a vehicle, other operations, such as changing of a wheel or filling the gas tank, require a correct localization of the different parts of the vehicle on which operations are to be performed. The paper describes a two-step approach to localize vehicle parts over the surface of a vehicle in front, rear and lateral views capitalizing on a novel approach based on bio-inspired visual attention. First, bounding-boxes are determined based on a model of human visual attention to roughly locate parts of interest. Second, the bounding-boxes are further searched to finely tune and better capture the boundaries of each vehicle part by means of active contour models. The proposed method obtains average bounding-box localization rates over 99.8% for different vehicle parts on a dataset of 120 vehicles belonging to sedan, SUV and wagon categories. Moreover, it allows, with the addition of the active contour models, for a more complete and accurate description of vehicle parts contours than other state-of-the-art solutions. This research work is contributing to the implementation of an automated industrial system for vehicle inspection.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

The increase in the number of vehicles on the roads generates new requirements for car dealers and garages to offer fast and efficient service. While more complicated operations within the car will continue to require the high expertise of human technicians, many simpler operations such as filling the gas tank or changing wheels could become automated and executed with the help of servicing robots in the near future. The successful execution of such operations requires an as-accurate-as-possible localization of specific vehicle parts to avoid excessive movement of the robotic equipment that is usually time-consuming and leads to safety concerns.

The work in this paper addresses this issue by proposing a novel solution to the problem of localization of vehicle parts such as wheels, windows, headlights and rear lamps, front and rear bumpers, lateral mirrors and gas tank trap in a set of images representing multiple views of vehicles. It initially proposes an original

bounding-box approach to roughly locate vehicle parts based on biological visual attention. Human visual capabilities are a rich source of inspiration for the improvement of computational vision algorithms, since human beings show a significantly superior performance in interpreting visual scenes and extracting regions of interest than most of the current machine vision technologies. Mimicking the role of human visual attention that extracts relevant regions of interest within a visual scene, a computational visual attention model is used in the context of this work to identify areas of interest over the surface of a vehicle. Visual attention models output a representation, called a saliency map (SM), in which areas of high visual interest are highlighted [1]. The projection curves on the two axes of the binary converted saliency map, which appears as a features-of-interest map, contain important information on the location of different parts of a vehicle and allow for the identification of a set of bounding-boxes that contain those vehicle parts. The set of bounding boxes is spatially adjusted over the vehicle surface according to the category of the vehicle. Active contour models (ACMs) [2] are then applied within the boxes to obtain a finer description of the contour of each part of interest. The proposed method goes therefore beyond state-of-the-art work on vehicle parts localization that is in general limited to fixed sized bounding boxes. It also allows the identification of a significantly larger number of parts in different views, compared to the current

\* Corresponding author at: Université du Québec en Outaouais, 101 Saint-Jean-Bosco, Gatineau, QC J8X 3X7, Canada. Tel.: +1 819 595 3900x1527.

E-mail addresses: [ana-maria.cretu@uqo.ca](mailto:ana-maria.cretu@uqo.ca) (A.-M. Cretu), [ppayeur@uottawa.ca](mailto:ppayeur@uottawa.ca) (P. Payeur), [laganier@uottawa.ca](mailto:laganier@uottawa.ca) (R. Laganière).

literature that is generally restricted to a limited number of vehicle parts and to a single view of a vehicle.

The paper is organized as follows: Section 2 presents the related work on the topic and Section 3 describes the proposed method for localization and fine tuning of vehicle parts. Section 4 discusses the experimental results and compares them with state-of-the-art solutions. Finally, the conclusions and the future work are presented in Section 5.

## 2. Literature review

While there has been a lot of interest in the detection of vehicles in images (a survey is available in [3]), there are relatively few papers dedicated to the localization of vehicle parts in the literature. Most of the existing work is concerned with the location of the vehicle license plates and logos [4–9]. In [4], an adaptive segmentation technique called Sliding Concentric Windows is employed to locate the license plate. Chacon and Zimmerman [5] apply Pulse Coupled Neural Networks (PCNN) to generate candidate regions that may contain a license plate. To solve the same problem, Guo et al. [6] propose a hybrid method based on PCNN and wavelet analysis. The method proposed in [7] to extract the vehicle plate region uses specific knowledge, such as the higher density of the plate region due to the presence of characters.

A three-layer neural network, trained with texture descriptors computed from the front image of vehicles is used for car plate and logo recognition in [8], while in [9] the vehicle license plate location is followed by a coarse-to-fine method to identify the logo based on a phase congruency feature map. A summary of various other techniques for plate localization and recognition is presented in [10].

A simplistic solution for side-view car fitting based on a sketch vehicle template is proposed in [11]. Brehar et al. [12], identify the pillars in side views of vehicles, based on a rough selection of objects that are likely to have one or two wheels based on circular symmetry, followed by an adaptive boosting classifier built using histograms of oriented gradient features. Because the solution is based on the detection of wheels, the approach cannot be directly applied for frontal or rear views. Lam et al. [13] identify different vehicle components such as: roof, windshield, bonnet, side windows, lower front of car (grille, headlight and front bumper) and lower side of car (wheels and door panels) in a monocular traffic image sequence using a topological structure of the vehicle based on multi-scale textural couriers. The vehicles are divided into multiscale regions based on the center of gravity of the foreground vehicle mask and the calibrated-camera parameters. A series of three key feature points, selected based on the assumption that cars have generally a windshield and headlights, allows for the identification of parts. The method is not directly extendable to rear views as the windshield is not visible and there is an uncertainty that the same key feature points would be useful in this case. Chang and Cho [14] detect in real-time the bumpers and wheels of a moving vehicle, viewed from the lateral side only, using Haar-like features. The algorithm capitalizes on temporal correspondence to reduce the search zones for parts, by verifying when a vehicle enters, and respectively exits the video frame. For this reason this approach cannot be directly applied for frontal or rear views. More recently, Chávez-Aragón et al. [15] proposed a method for vision-based detection of vehicle parts such as bumpers, windows, door handles, wheels, lateral mirror, windshield, center, roof, headlights and rear lamps in lateral views. The approach is using a geometrical model to determine feasible search areas for parts and a cascade of boosted classifiers based on Haar-like features to detect the parts within each feasible zone, in a fixed sized bounding box style. The algorithm first identifies the two wheels in a

side view of a vehicle, and their relative position in order to determine the location of other parts with respect to it. Therefore this solution is not suitable neither for the identification of parts in frontal or rear views. In [16,17], the license plate and rear lamps (only red ones), are localized in rear images of cars using their distinctive color, texture, and expected regions in the context of an urban traffic surveillance application. This sort of approach does not work on the detection of parts that are not clearly identified as having a different color or texture or on different views of a vehicle.

The current paper builds on previous work of the authors on the topic [18,19] and goes beyond state-of-the-art solutions by initially proposing a novel, better performing bounding box approach based on visual attention to roughly identify the position of vehicle parts in a first phase. Following a simple initialization stage, in which the user selects from a series of bounding boxes over the surface of a single vehicle the ones of interest for his/her application, the novelty of the approach consists in adapting a visual attention model to automatically adjust the bounding boxes to better fit these parts of interest for any other vehicle category. Therefore this initialization step allows for a smooth and simple adaptation to any specific application that requires the identification of vehicle parts and contributes to ameliorate the performance by only processing relevant information. A significant advantage of the proposed solution is that, unlike other approaches available in the state-of-the-art literature, it can localize a larger number of parts and operates from different views (e.g. side, frontal or rear) of a vehicle. The solution is also further improved in order to obtain a finer description of the contour of each part discovered in a given view by using ACMs in each bounding box. A thorough comparison performed with similar work shows the superior performance of the proposed approach.

## 3. Localization of vehicle parts based on visual attention and active contour models

The proposed system for localization of vehicle parts is composed of two major steps: an initialization step and a refinement step and its main blocks are illustrated in Fig. 1.

In the initialization step, all the images in the dataset are aligned to a reference image (usually the one belonging to the first vehicle in the dataset) and their saliency maps (SMs) are built based on a model of human visual attention. The category of each vehicle is determined based on the SM using the solution proposed in [19]. For each category of vehicles (here sedan, SUV and wagon categories are considered), an average SM model,  $SM_{avg.cat.view}$ ,  $cat \in \{sed, SUV, wag\}$ ,  $view \in \{1,2,3,4\}$  is built for each view of a vehicle by summing the individual SM models viewed from a given direction and dividing the resulting model by the number of vehicles within the category, or in other words, by the number of individual SM models. The views for each vehicle are provided by four distinct cameras situated around the vehicle, one in front (view 1), one in the back (view 2) and the other two on the lateral sides (views 3 and 4). Due to this specific setup that is compatible with the proposed application, the view from which the vehicle is seen in an image is known because it comes from a given camera.

This average SM model serves as a basis for the identification of category-specific bounding boxes  $BB_{cat.view}$ . The bounding boxes are determined by projecting the average SM model on its X and Y axes and extracting the local minima and maxima over the projection curves. These local extrema contain important information on the position of different parts of interest, as will be detailed in Section 3.1. The coordinates of the local minima and maxima serve as coordinates for the bounding boxes. The average model of

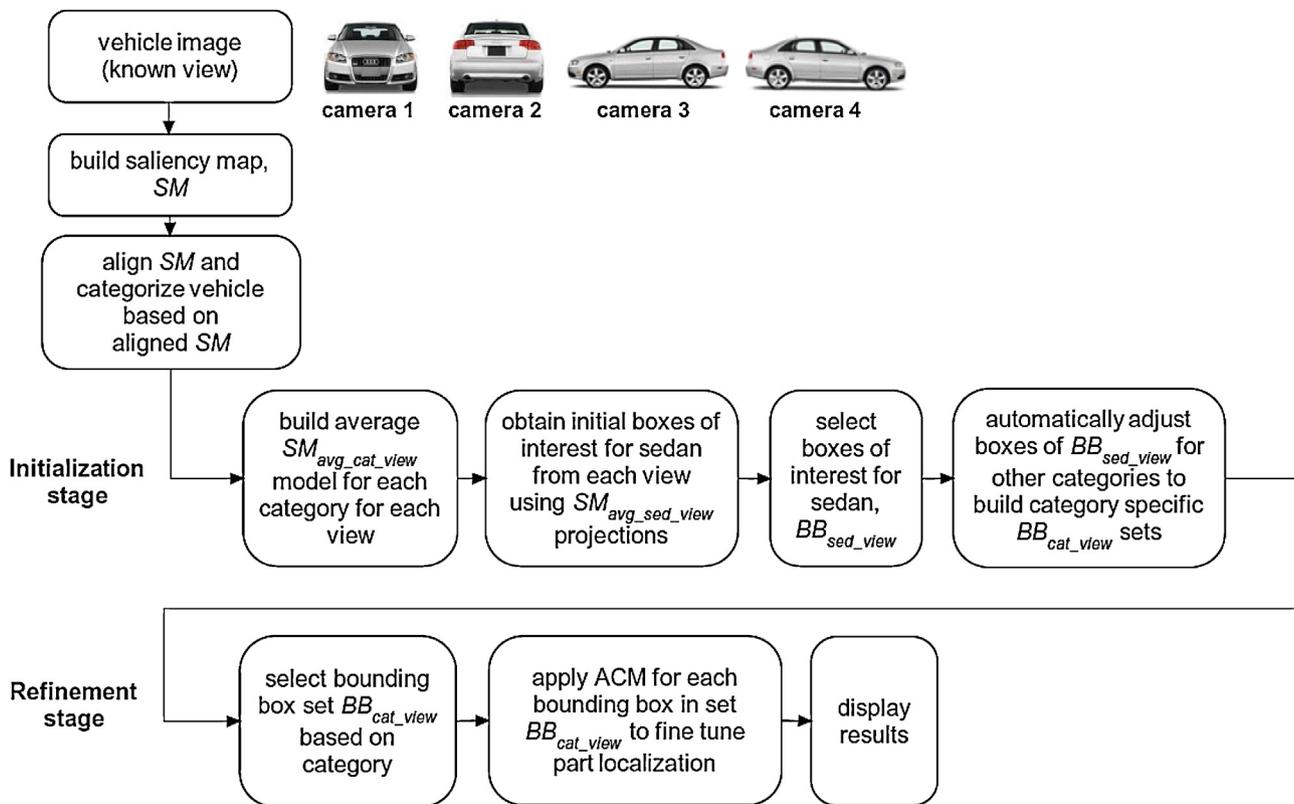


Fig. 1. Overview of the proposed system.

a sedan,  $SM_{avg\_sed\_view}$ , is considered as a reference vehicle model and is the only one for which the user selects among the boxes, the ones that contain parts of interest. Therefore the proposed solution requires the user to only intervene in the initialization stage by choosing the bounding boxes in one vehicle saliency map from a desired viewpoint. In the case illustrated in Fig. 1, the choice of bounding boxes is made over the average sedan saliency map (the user intervention takes place in the third box of the initialization stage in Fig. 1), but any other vehicle category could be used as well without impact on the performance. This initial setup can be beneficial in the sense that the user might be interested in specific vehicle parts and/or in a specific viewpoint for a particular application. The selection of only the required parts of interest and viewpoints can have a positive impact on the associated computation time, as it will avoid the non-necessary computation for all parts and all viewpoints. This model is then automatically adjusted to fit the position of the bounding boxes for other vehicle categories, as further detailed in Section 3.2. By computing the difference between the local minima and maxima on the X axis obtained on the average SM of any other category and the local minima and maxima on the X axis of the sedan, the boxes can be displaced automatically to the left or right to better fit the shape characteristics of the new category. In a similar way, the boxes can be moved up or down based on the local minima and maxima on the Y axis.

Each time an image of a vehicle is presented to the system, the SM is computed and this information is used to categorize the vehicle. The same information is also used to align the image with the reference image. Up to this stage, the proposed method identifies areas where a part could be situated and does not attempt to precisely locate the part within its corresponding bounding box. To further refine the contours of each part of interest, the solution makes use of active contour models (ACM) within each adjusted

bounding box. One ACM is used for each part of interest. The ACM is initialized within the bounding box corresponding to the part of interest as determined in the initialization step, and is applied on the initial color image. The advantage of using ACMs is that, even if the part of interest is partially out of the determined bounding box, by iteratively moving towards the contour of the part of interest, the model adapts to fully capture it. The procedure is detailed in Section 3.3.

Finally, the resulting contours are overlaid on the initial image for display and performance evaluation. This information can be further used by a robot to efficiently and accurately reach any part of interest over the vehicle to perform automated servicing provided the cameras are properly calibrated within the robot's workspace.

### 3.1. Bounding box localization of vehicle parts based on a model of human visual attention

The main idea behind the bottom-up visual attention computational model of Itti et al. [1] that is largely followed in this work, is to compute several features derived from a color image provided as input and fuse their saliencies into a representation called saliency map. Several features such as: the intensity ( $I = (R + G + B)/3$  where  $R$ ,  $G$  and  $B$  are the red, green and blue color channels), the color (represented by the red-green,  $RG$ , and blue-yellow,  $BY$ , color opponency) and the orientation are computed in parallel. Feature-dependent saliencies are then calculated for each of the three feature channels. Center-surround operations modeled as a difference between fine and coarse scales are applied on all features. Each feature set is stored in feature dependent saliency maps, so called conspicuity maps, denoted  $\bar{C}$  for color,  $\bar{O}$  for orientation and  $\bar{I}$  for intensity, in form of grayscale images where the intensity of each pixel is proportional to its saliency. The color

**Table 1**  
Comparison of average recognition rates for the three vehicle categories.

	Sedan (%)	SUV (%)	Wagon (%)
SIFT key points	73.2	67.9	81.2
Harris corners	76.5	67.9	81.0
DoG features	88.2	89.6	93.5
Gabor features	87.1	92.4	93.5
SM-based features	91.2	94.4	94.8

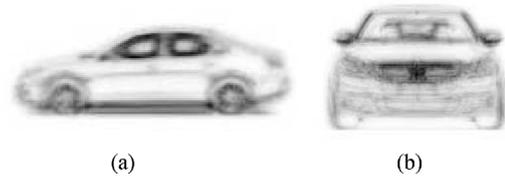
conspicuity map is calculated as  $\tilde{C} = \oplus_{c,s} N(RG(c,s)) - N(BY(c,s))$ ,  $c \in \{2, 3, 4\}$ ,  $s = c + \delta$ ,  $\delta \in \{3, 4\}$ , where  $c$  represents the center-scales,  $s$  are the surround-scales,  $\oplus$  represents an across-scale addition,  $N$  represents the normalization operation by  $(G - L)^2$ , where  $G$  is global maximum in the map and  $L$  the average of local maxima in the same map. The color opponency components  $RG(c,s)$  and  $BY(c,s)$  are calculated as  $RG(c,s) = |(R(c) - G(c)) \ominus G(s) - R(s)|$  and  $BY(c,s) = |(B(c) - Y(c)) \ominus Y(s) - B(s)|$  respectively, where  $\ominus$  represents an across-scale subtraction operation, as in [1]. The intensity conspicuity map is computed as:  $\tilde{I} = \oplus_{c,s} N(|I(c)| \ominus |I(s)|)$ .

Finally the local orientation information is obtained from the intensity image  $I$ , using oriented Gabor pyramids of different scales and different preferred orientations  $\theta$  and the orientation conspicuity map is  $\tilde{O} = \sum_{\theta} N(\oplus_{c,s} N(O_{\theta}(c,s)))$ , with  $O_{\theta}(c,s) = |(O_{\theta}(c) \ominus O_{\theta}(s))|$ ,

and preferred orientations  $\theta = \{0^{\circ}, 45^{\circ}, 135^{\circ}, 180^{\circ}\}$ . These three conspicuity maps for color, intensity and orientation are summed up in the final saliency map, using the formula  $SM = (\tilde{C} + \tilde{I} + \tilde{O})/3$ . The full implementation details are available in [1]. In this work, the computational attention model is used to identify the areas of interest in the images representing front, back and lateral views of vehicles belonging to the three categories, i.e. sedan, SUV, and wagon.

The saliency map  $SM$  plays a triple role in the context of this work. It first allows the identification of the category of vehicle. A method that was proposed for image-based vehicle classification based on SM and support vector machines is used for this purpose. Full details are available in [18]. The method achieves classification rates between 86.9% and 96.2% for the three categories of vehicles used in this paper when the views from front, back and lateral sides are considered separately and an average rate of 94.3% when the decision from all four views is considered. Additional tests were performed to compare the average recognition rates for the three vehicle categories and the results are shown in Table 1. It can be observed that the SM-derived features obtain better results for vehicle classification (with an average of 93.5% over the three above-mentioned categories) than SIFT key points (average of 74.1%), Harris corners (average of 75.1%), Difference-of-Gaussians (DoG) features (average of 90.4%) and Gabor features (average of 91%). It is important to state as well that the computation time of SM-based features is in line with the one of Gabor and DoG features.

Secondly, the SM provides the necessary information for the alignment of vehicles with the reference image. Images are aligned such that the center of all vehicles corresponds with the center of the vehicle in the reference image. The center is computed based on the width and height projection curves of the corresponding SM. In order to compute the width and height, the image representing the SM is initially converted to black and white, using the Otsu thresholding method to obtain  $SM_{bw}$ . The vertical and horizontal projection curves are built by summing all the columns of  $SM_{bw}$  to obtain width information,  $w$ , and all the rows to obtain height information,  $h$ . The horizontal projection curve is then searched starting from the left until a value different from 0 is identified. Each time a value of 0 (empty background) is encountered, the



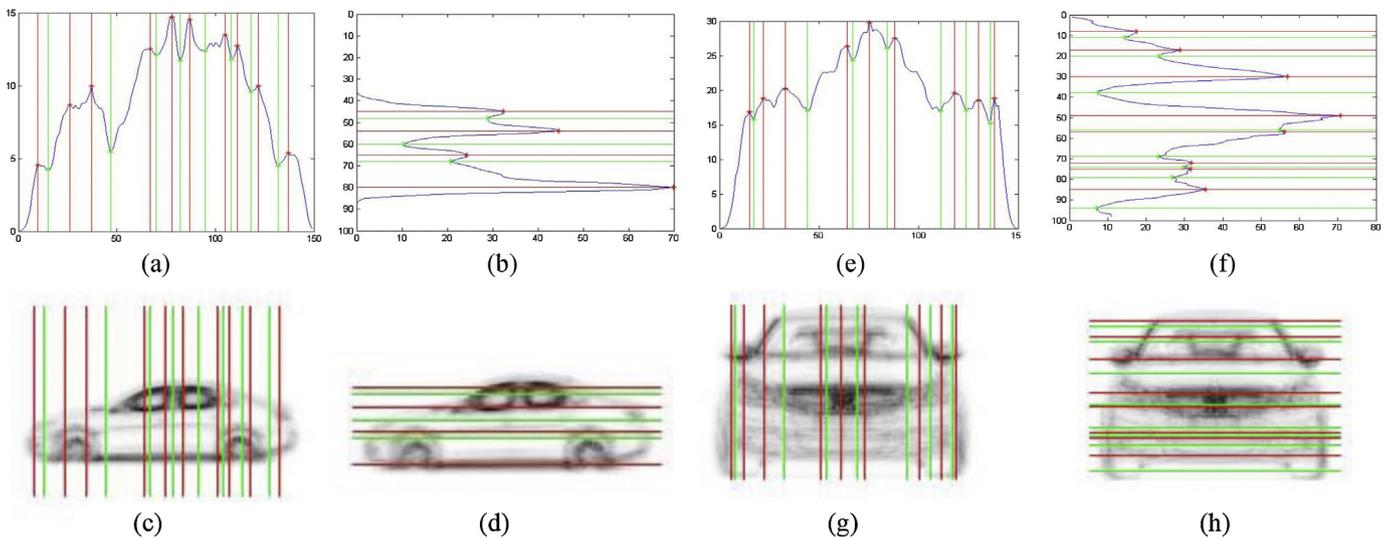
**Fig. 2.** Average saliency map for sedan category from (a) lateral and (b) front view.

value of the width,  $w$ , originally initialized to the full width of the image, is decreased by 1 pixel. When the first value different from 0 is encountered, the search from the left direction is stopped. The same procedure is used from the right direction by decreasing the remaining width value,  $w$ , until a value different from 0 is found. A similar top and down search is performed on the vertical projection curve to compute the height,  $h$ , of the vehicle in the image. The center of the vehicle is computed based on the width and height information, as:  $C(x,y)|_x = X_{nz} + w/2$ ,  $y = Y_{nz} + h/2$  or in other words, having the X coordinate equal to the coordinate where the first non-zero value appeared in the horizontal projection curve (where the body of the car starts),  $X_{nz}$ , plus half of the computed width, and the Y coordinate as the Y coordinate where the first non-zero value appeared in the vertical projection,  $Y_{nz}$ , plus half of the computed height. The reference center value is computed for the reference image. A translation is performed to move the center of any other image to the reference center and therefore align all vehicles. While the approach can be sensitive to scaling effects, it is constrained by defining a specific setup where the vehicle is parked in the bounded perimeter of the four cameras and therefore no significant scaling effects take place. Moreover, to further circumvent the problem, all images and the corresponding saliency maps can be rescaled to the size of the reference vehicle after the categorization stage (where the size plays significant role for distinction between vehicle categories).

Thirdly, the SM provides the necessary information for identification of vehicle parts. Fig. 2 shows the average SM model for the sedan class, computed over 41 samples of sedan in the dataset viewed from the lateral and front views.

The SMs are presented as negatives (e.g.  $1 - SM$ ) for better visualization of the results. In the image, the areas of higher interest are marked by darker shades. The projections on the X and Y axes of these average models are illustrated in Fig. 3a,b for the lateral view and in Fig. 3e,f for the front view, respectively. The local maxima and minima are not local maxima and minima from a strict mathematical sense, but rather peaks (local highest point, around which there are points lower by a given amount, e.g. 0.35 in the current application, on both sides) and valleys of the curves. Local maxima are displayed with red stars/bars and local minima with green in Fig. 3. Local maxima,  $M$ , and minima,  $m$  are used together to achieve better localization of different parts of interest along the projection profiles. Fig. 3c,g show the correspondence between the projection curves on the X axis and different parts of the vehicle viewed from lateral and front, with red vertical lines representing the values of local maxima and the green lines those of the local minima. One can notice, for example, in Fig. 3c that the first maximum corresponds roughly to the beginning of the headlight, while the second and third maxima frame roughly the location of the wheel. Similarly, local maxima along the Y axis, illustrated in Fig. 3d,h for the two views, provide an estimate of the location of parts. For example, in the lateral view in Fig. 3d, the front and back bumpers and the wheels are situated between the first and second local maxima (shown in red), starting from the bottom of the figure.

By using jointly the information on the two axes, as illustrated in Fig. 4a,c for the lateral and front view respectively, a grid is obtained that provides an approximate localization of the different parts of interest.



**Fig. 3.** Projections on X axis (a) for lateral view and (e) for front view, projections on Y axis (b) for lateral view and (f) for front view, with local maxima shown in red and local minima in green; and correspondence between projections on (c) and (g) the X axis, and (d) and (h) the Y axis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 4b,d shows the selected bounding boxes corresponding to the parts of interest. For example, in the lateral view illustrated in Fig. 4a, the front bumper is situated in the box that ends with the first local minima (green) on the X axis and between the two first local maxima (in red) on the Y axis from the bottom of the figure. The corresponding rectangle is shown in yellow in Fig. 4b, which also displays all the parts of interest visible in the lateral view: the bumpers are located within yellow rectangles, the head and back lamps in green, the wheels in red, the door handles in magenta, the windows in blue, the mirror in cyan and the gas trap within the orange rectangle. For the front view, the same colors are used to localize the wheels, mirrors, windshield and front lamp. In this case, the grille is marked by a magenta box and the vehicle logo by a yellow box. In a similar way, average models can be built for other vehicle categories. The discovery of parts of interest from the grid is solely based on the projection curves. In the current solution, their position is user initialized for the sedan class and adjusted automatically for all other vehicle categories. If the user desires to add other parts of interest, he can choose manually the corresponding bounding boxes for these parts in the grid (only once and only for the sedan class).

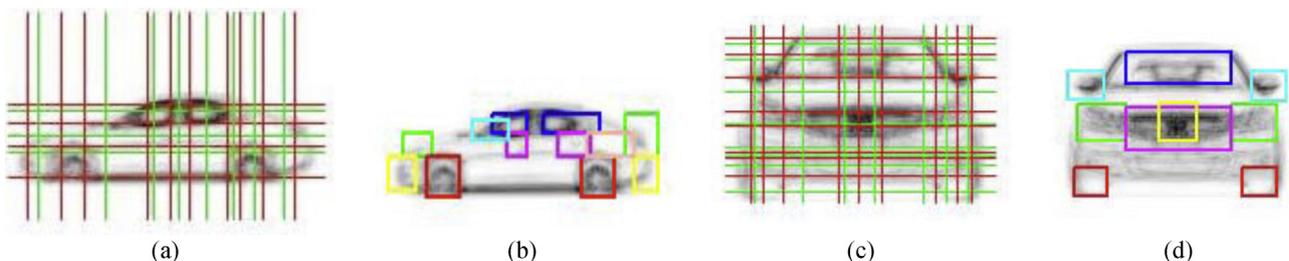
Additional tests were conducted in order to evaluate the capability of classical edge detectors to perform the same task and to potentially replace the SM-based solution for rough identification of parts of interest. In particular, Canny and Sobel algorithms were used in a similar manner to the SM-based approach, to initially build an average edge model, corresponding to the average saliency map illustrated in Fig. 2, and the resulting curves were projected on the X and Y axes as in Fig. 3a,b and e,f to find associations between extrema and the parts of interest. No association could be identified

in this case, with or without the application of filtering operation to smooth out the curves, due to the extremely high number of extrema and the roughness of the curves.

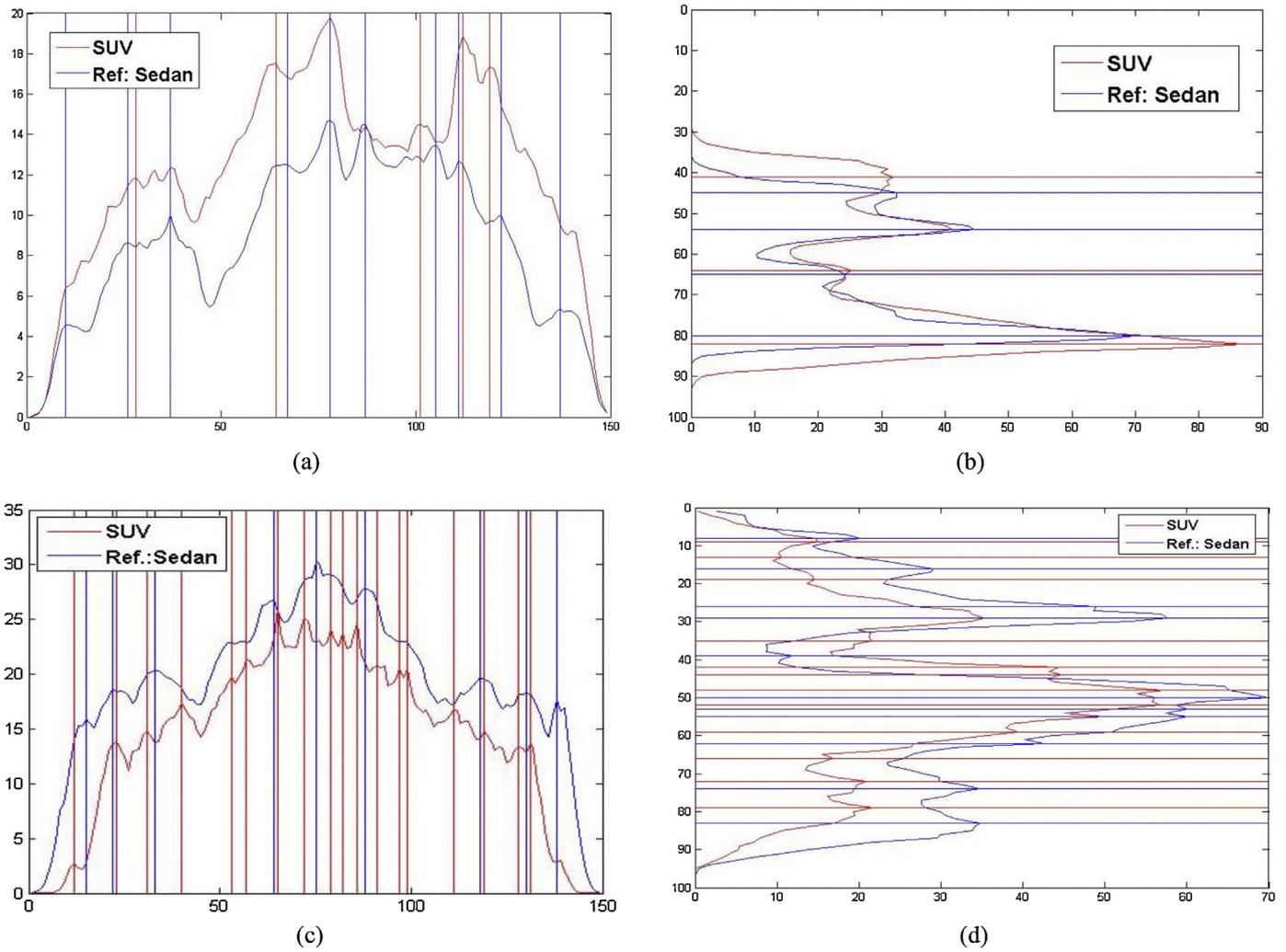
### 3.2. Adjustments of the location of bounding boxes for other vehicle categories

The knowledge available about the local maxima and minima of the projection curves on the two axes allows for the automated adjustment of the position of bounding boxes to better fit the average models from different views of other categories of vehicles. Fig. 5 shows the difference between the projection curves of the SUV average model (in red), computed over 49 samples of SUVs in the dataset, and those of sedan (in blue) for two views. Only the maxima are shown not to overload the figures. The first, third and last maxima on the X axis in Fig. 5a as well as the third maxima from the bottom in Fig. 5b correspond for the two vehicles, but are not visible in the figure due to their overlap.

One can notice that some small adjustments can be made to fit the bounding boxes of the sedan over the SUV model. For example, to better fit for the SUV category, the front wheel bounding box, situated between the second and third maxima on the X axis in Fig. 5a and between the first and second maxima from the bottom on the Y axis in Fig. 5b, has to be moved slightly to the right (proportional to the difference between the second local maxima for sedan and second local maxima for SUV) and slightly lower (proportional to the difference between the first local maxima for sedan on the Y axis and the first local maxima for SUV on the same axis). In those cases where more peaks are present in the profile of another category, the average of their values is computed and mapped to the



**Fig. 4.** Grid for definition of bounding boxes for (a) lateral view and (c) front view, and bounding boxes corresponding to parts of interest in (b) lateral view and (d) front view. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Differences between SUV (in red) and sedan (in blue), first column: X profile for (a) side view and (c) front view, and second column: Y profile for (b) side view and (d) front view. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

closest value in the sedan profile. The algorithm for the adjustment of boxes based on local maxima,  $M_Y$ , along the Y direction is summarized below in pseudo-code, where  $M_{ref,Y}$  denotes the local maxima on the Y axis derived from the reference model (e.g.  $SM_{avg, sed.view}$ ).

#### Adjustment algorithm for bounding boxes

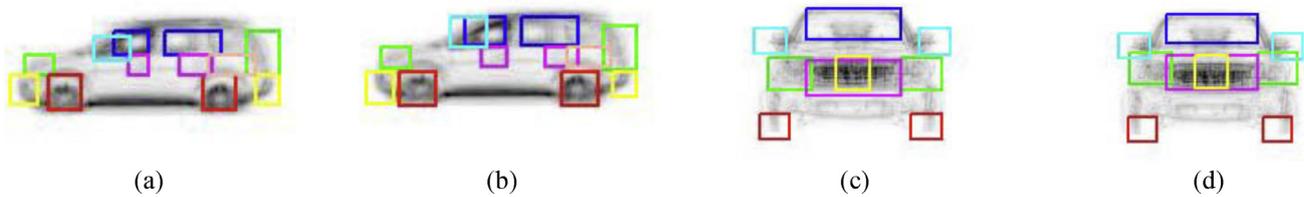
```

Set the amount of adjustment,  $adj_{MY} = 0$ ;  $amnt = 5$  (number of pixels to the left
and right of the local maxima in the reference model where a search is
performed for local extrema in the profile of the category to be adjusted);
for all the elements of  $M_{ref,Y}$ 
  for all the elements of  $M_Y$ 
    build a list of values in  $M_Y$  within the  $amnt$  number of pixels:
     $M_{YL} = \{M_Y | M_{ref,Y} - amnt < M_Y < M_{ref,Y} + amnt\}$ ;
  end
  if  $M_{YL} = \emptyset$  ( $M_{YL}$  is empty) then  $adj_{MY} = 0$ ;
  else
    if  $M_{YL}$  has one element
      if the local maxima coincide in the two profiles  $M_{YL}(1) = M_{ref,Y}$  then
         $adj_{MY} = 0$ 
      else the adjustment is the amount of difference between the local maxima
         $adj_{MY} = M_{YL}(1) - M_{ref,Y}$ ;
      else
        sort the elements of  $M_{YL}$  in descending order to build  $M_{YL.sort}$ ;
        amount of adjustment is the difference between the average of values in
        the profile to be adjusted minus the value of the reference maxima:
         $adj_{MY} = \sum[(M_{YL.sort}/size(M_{YL.sort})) - M_{ref,Y}]$ 
      end
    end
  end
end
end

```

The main idea of the algorithm is to search for similar local extrema over the X and Y axes in the projection profile of the reference average saliency map of a vehicle (sedan in this case, e.g.  $SM_{avg, sed.view}$ , but any other category could be chosen as well) and the projection profiles of any average saliency map belonging to other vehicle categories. The search takes place in the projection profile of the average saliency map of the new category in the areas corresponding to each local extrema of the projection profile of the reference saliency map over the X and Y axes respectively, and within  $amnt$  pixels to the left and right of each extremum. Once an extremum is identified in the projection profiles of the new vehicle category, the amount of displacement is calculated between the local reference extremum and the local extremum of the new category projection (e.g.  $adj_{MY}$  for the Y axis and similar for the X axis). The corresponding bounding box, situated at the local extrema in the reference model is then moved in space with this displacement over X and Y axes respectively, resulting in a better fit of the bounding box for the new vehicle category. If no extremum is identified in the search region or the displacement is 0, the bounding box keeps its original position.

Only those local maxima are considered that frame the selected bounding boxes representing different parts, namely the ones that determine the boxes illustrated in Fig. 4b or d. A similar procedure is followed for the local maxima on the X axis and the local minima on the X and Y axes.



**Fig. 6.** Reference bounding boxes of sedan superimposed over the average SUV model for (a) lateral view and (c) front view, and adjusted position of bounding boxes for SUV for (b) lateral and (d) front view.

**Table 2**

Parameters for ACM for different vehicle parts.

Vehicle part	Pixels around centre on X axis/pixels around centre on Y axis/number of iterations/ $\mu$		
	Front view	Rear view	Side views
Wheels	$\pm 5/\pm 5/20/0.05$	$\pm 5/\pm 5/20/0.05$	$\pm 2/\pm 2/20/0.05$
Windows	$\pm 20/\pm 5/25/0.05$	$\pm 20/\pm 5/25/0.05$	$\pm 4/\pm 3/50/0.05$
Bumpers	N/A*	N/A	$\pm 3/\pm 3/50/0.05$
Handles	N/A	N/A	$\pm 2/\pm 2/40/0.05$
Lights	$\pm 5/\pm 5/40/0.05$	$\pm 2/\pm 2/40/0.05$	$\pm 2/\pm 2/40/0.05$
Grille	$\pm 4/\pm 10/95/0.05$	N/A	N/A
Mirror	$\pm 4/\pm 4/15/0.05$	N/A	$\pm 2/\pm 2/15/0.05$
Logo	$\pm 2/\pm 1/45/0.05$	$\pm 2/\pm 1/45/0.05$	N/A
Gas trap	N/A	N/A	$\pm 1/\pm 5/20/0.05$

\* N/A denotes parts that are not perceived in the specific view.

Fig. 6a,c illustrates the uncorrected bounding boxes from for the sedan model that are directly projected from the SM to the SUV. Fig. 6b,d shows the better fit of bounding boxes when adjustments are performed. It can be seen that they fit better the different parts of interest. For example, when comparing Fig. 6a with b, the corresponding bounding boxes for the wheels, in red, and for the lateral windows, in blue, are better centered over the wheels and windows surfaces respectively and they cover the wheels and windows entirely in Fig. 6b, which is not the case in Fig. 6a. The same states for the headlamps in Fig. 6d with respect to those in Fig. 6c. Adjustments for any other category of vehicles are obtained in a similar way.

### 3.3. Fine boundary tuning for vehicle parts using active contour models

In order to refine the boundaries of the different parts of interest for the vehicle, active contour models (ACMs), particularly the active contour models without edges of Chan and Vese [2] are applied over the initial color image, as briefly described in the beginning of this section. The ACM is initialized for all parts of interest with a small rectangle situated around the center of the bounding box. The coordinates of the center of the box ( $x_c, y_c$ ) are computed based on the width and height information, as described in Section 3.1, and a given number of pixels is added to the left and right (along X axis), up and down (along Y axis) of this center to create the small initialization rectangle for the ACM. In order to determine the number of pixels to be added to the center for each of the parts of interest, the average size (width along x axis, denoted

$w_a$  and height, along y axis, denoted  $h_a$ ) of each bounding box corresponding to this part is calculated over all vehicle categories, namely sedan, SUV and wagon. The coordinates of the initialization rectangle are then computed as  $x = x_c \pm \text{int}(w_a/4)$ ,  $y = y_c \pm \text{int}(h_a/4)$  and are illustrated for each part of interest in Table 2. Given that these parameters are automatically calculated, the user does not need to select them empirically neither for the dataset in the context of this work, nor for any other dataset. One can notice that the size of the initialization rectangle is different for the front/rear and the lateral views and this occurs due to the change in the size of different parts with respect to the entire image. It is also worth mentioning that in the dataset used for experimentation, the size of vehicles viewed from the sides is relatively smaller with respect to the front and rear views and this also impacts the size of the rectangle used for ACM initialization. The number of iteration steps is determined experimentally as the minimum number of steps required to capture a part of interest and is kept constant for each part for all the vehicles. Aside from the parameters shown in Table 2, some additional specific constraints are imposed for different parts as it will be further discussed.

Fig. 7 illustrates an example for the detection of the front window for the side view of a sedan. Fig. 7a shows the ACM model, Fig. 7b illustrates the result after the holes in the ACM model are filled. It can be observed that (as in almost all cases dealing with internal parts) the ACM captures more than the desired part.

Therefore only the part of ACM situated within the initial bounding box (including a small tolerance of up to 3 pixels, to cope with small errors in the bounding-box positioning) is considered, as illustrated in Fig. 7c. The boundaries of the detected regions of



**Fig. 7.** (a) ACM model, (b) ACM model after hole filling, (c) model constrained in the bounding box corresponding to front window and (d) ACM contour superposed over initial image.



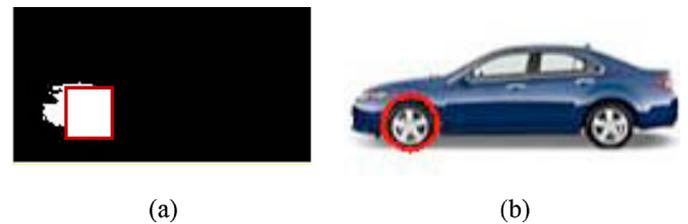
**Fig. 8.** Detection of (a) red and (b) non-red rear lamp. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

interest are displayed to frame the fine localization of different parts of interest over the initial image in Fig. 7d. The same approach is used to detect the front and rear bumper, the grille, the headlight and the door handles. For the handles, an additional constraint referring to the size of the part (e.g. larger than 3 pixels) is imposed to allow for the elimination of small outliers in the ACM model that result from the changes in door curvature close to the handles and from the close proximity of the door seam. Similarly, maximum area constraints are imposed on all small sized parts like the mirror and logo to eliminate areas of the vehicle that are wrongly included due to their proximity to the part of interest in the ACM model.

As the rear lamp is usually red, it has been noticed during the experimentation that the addition of the color information is helpful to better define the contour of the rear lamp. In order to detect red color in the image, the grayscale image is subtracted from the red channel of the color image of a vehicle and the result is thresholded to obtain a black and white image in which the red areas are marked with white and all the other areas with black. Only those areas of red color situated within the bounding-box for the rear lamp are considered. An AND operation is performed between the red areas and the result of the ACM segmentation constrained in the window of interest to detect the contour of the rear lamp. In some cases, the color of the rear lamp is not red, as in Fig. 8c (in which case there are no white regions in the thresholded image) or the car is red (almost all the thresholded image is white). In these cases, only the results of ACM are considered for the fine localization of the rear lamp. Fig. 8b illustrates the successful detection of a red rear lamp for the vehicle in Fig. 8a, while Fig. 8d shows the successful detection of a rear lamp that is not red, for the vehicle illustrated in Fig. 8c.

In order to detect the boundaries of the wheels, a circular Hough transform is applied over the image transformed to grayscale. In order to compute the approximate radius of circles to be found by the transform and therefore shorten the computation time, an ACM is first applied in the bounding-boxes corresponding to the two wheels marked by red rectangles in Fig. 4b. The small empty holes in the ACM model that result due to the threads in the wheel are filled and the area of the resulting model is computed, again restricted within the corresponding bounding box as shown in Fig. 9a.

This area provides an estimation of the radius of each wheel, as being equal to the square root of the area over  $\pi$ . This radius ( $\pm 3$  pixels) is used as a guide for the Hough transform. As well, the centers of the circles to be identified by the Hough transform are constrained within the predetermined bounding-boxes to eliminate false detections. An example of correct wheel detection is



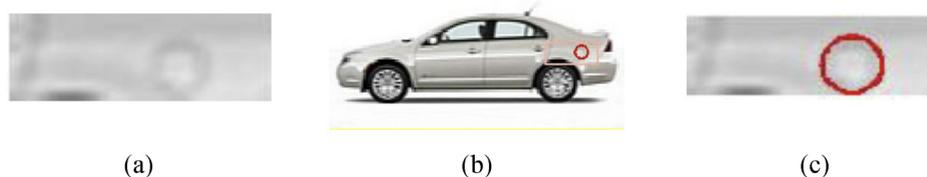
**Fig. 9.** (a) ACM model for front wheel and corresponding bounding box, and (b) wheel identification with circular Hough transform.

illustrated in Fig. 9b. A similar procedure is applied for the identification of the gas tank trap, since it is in most cases circular. The circular Hough transform is applied in this case within the area of interest over the top of the back wheel, marked by the orange rectangle and shown enlarged in Fig. 10a,b.

Since the dimension of the gas trap is unknown, the possible circle radiuses are set in a rather large interval between 7 and 15 pixels. An example of the identified circle is illustrated in Fig. 10b. The identified circle, if found, is translated back to the original coordinates (coordinates of the center plus the coordinates of the left bottom corner of the window of interest) within the bounding box in the original image, as shown in Fig. 10c. Since most of the gas traps are round in our dataset, the performance of the circular Hough transform was found to be satisfactory. The identification of other shapes of gas tank traps using ACMs is complicated due to the proximity of the door seam, of the fender and of the rear lamp, and due to the wide space over which the trap can be located. Alternative solutions will be further studied to better address this issue.

#### 4. Experimental results

A set of 120 vehicles [20] belonging to the sedan, SUV and wagon categories, is used for experimentation. Similar performance was obtained over 180 vehicles including also sport car and pickup truck categories. For each of the vehicles, 4 views are available as illustrated in Fig. 1: front and back views and the two lateral views. Each image in the dataset has  $99 \times 155$  pixels. The vehicles in the dataset are presented against simple white background. As the proposed solution is meant to be incorporated in a vehicle inspection application to operate in a garage, the background is known in this context and can be subtracted prior to the application of the algorithm. However, the solution has also been tested on images with various



**Fig. 10.** (a) Bounding-box, and (b), (c) successful localization of gas trap.

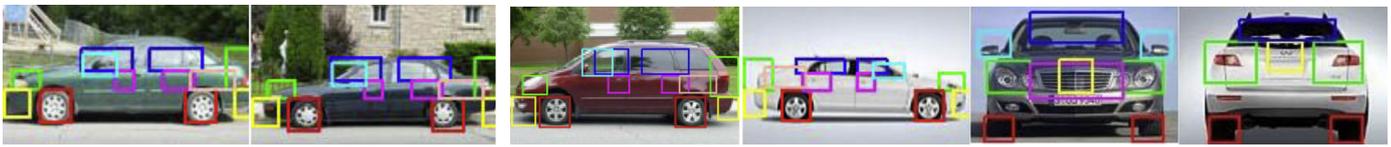


Fig. 11. Examples of parts located by bounding boxes on test images with various backgrounds.

backgrounds, as shown in Fig. 11, and demonstrated similar performance if the background is relatively uniform in colour, therefore not capturing the visual attention on parts that do not belong to the vehicle. The performance remains stable as long as there is no significant scaling between the vehicle in the test image and the ones in the datasets, as it was previously discussed in Section 3.1.

Fig. 12 shows several examples of bounding boxes and the corresponding ACMs obtained using the proposed method for the three categories of vehicles in the dataset and for the front, rear and lateral views. The performance for the passenger lateral view is very similar to the driver side and is therefore not illustrated. One can notice that both the bounding boxes and the ACMs are, with few exceptions, very well localized over the surface of the vehicles. It is worth mentioning that the bounding boxes indicate candidate regions where parts would be situated without trying to detect them in the image. For example, the gas trap appears in the side views in column 5 of Fig. 12, even if no gas trap is present on that side of the vehicle. While the fine tuning with ACM is not always perfect as it can be noticed in the last column of Fig. 12 (e.g. different shape of door handles in second line, partially covered front handle in 7th line, missed back handle due to unnatural placement in 10th line, or smaller windows in 9th line and 11th line), it provides in most cases better results in terms of localization than the simple bounding-box approach.

4.1. Performance evaluation

In order to quantify the accuracy of the proposed approach, a quantitative measure is computed as the percentage of the surface of the actual vehicle part covered by its corresponding bounding box and by its ACM model, respectively. The ground truth surface of the actual vehicle part is obtained by manual segmentation in a rectangular bounding-box style such that the box has the smallest size possible that covers completely the part.

Similarly, the bounding box that corresponds to the ACM model is calculated as the largest rectangular box that contains the entire contour. At maximum, this box can have the same size as the corresponding bounding box within which the ACM is initialized, as explained in Section 3.3. This occurs in the case when the part is not identified by the ACM. Tables 3–6 summarize the results obtained for each category of vehicle and for each of the views.

Table 3 Accuracy of bounding box (BB) and ACMs localization for side view.

Vehicle Part	Percentage of area (%) in BB/ACM			Parts found (%) BB/ACM
	Sedan	SUV	Wagon	
Front wheel	88/100	92.8/100	84.8/100	100/100
Rear wheel	89.9/100	90.5/100	90.7/100	100/100
Front window	95.9/90.6	94/95.3	100/89.61	100/100
Back window	96.7/93.8	95.5/94.6	95.4/91.3	100/100
Front bumper	100/97.3	100/100	100/100	100/99.1
Rear bumper	100/92.2	100/93.3	100/93.6	100/92.9
Front handle	95.3/80.9	95.5/90.6	96/78.5	95.6/89.2
Rear handle	95.2/84.7	99/87	99/73.1	97.6/93.6
Head-light	96.5/69.3	100/70	96.7/76.2	100/89.3
Rear light	99.8/91.8	99.5/88.4	98.1/82.1	100/98.8
Mirror	100/70.4	100/90.6	100/87.5	100/87.6
Gas trap	100/89.4	100/98.5	99/98.3	100/94.4
Average	96.4/88.4	97.2/92.4	96.6/89.2	99.4/95.4

Table 4 Accuracy of bounding box (BB) and ACMs localization for front view.

Vehicle Part	Percentage of area (%) in BB/ACM			Parts found (%) (BB/ACM)
	Sedan	SUV	Wagon	
Left wheel	92.9/100	92.5/95.1	95.9/97.5	100/100
Right wheel	92.8/100	93.2/95.6	92.3/98.3	100/100
Wind-shield	99.4/78.4	90.8/83.6	90/78.6	100/100
Grille	99.9/98.3	100/99.6	98.3/94.8	100/100
Logo	100/84.3	100/86.8	100/78.8	100/86.2
Headlight left	93.1/74.4	94/85.8	98.5/78.4	100/96
Headlight right	93.3/81.4	89.2/83.4	91.6/72.2	100/95.1
Mirror left	94.4/63.5	98.6/68.1	100/73.6	100/85
Mirror right	94.4/68.6	94.4/67.6	100/71.3	100/86.6
Average	95.6/83.2	94.7/85.1	93.3/82.6	100/94.3

In order to enable the comparison with other solutions proposed in the literature, which report the percentage of cases when a part is found, the fifth column in Tables 3–5 reports the percentage of cases when a part is found within its bounding box and within its corresponding ACM respectively. A vehicle part is considered found when at least 45% of its surface is within its corresponding bounding box. From Tables 3–5 it can be observed that there is no significant difference between the performance for sedan, SUV and wagon (within 3%) neither for the bounding boxes, nor for the ACM. The overall performance deteriorates slightly as result of ACM application for all categories, both in terms of surface of vehicle part covered (from an average of 94.4% for bounding box to 86.6% for ACM) and in terms of parts found (from an average of 99.8% for bounding box to 95.86% for ACM). It is worth mentioning nevertheless that this is partially explained by the fact that the bounding box is much larger than the real part, while the size of ACM is usually much closer to the real size of the part, and ACM also locates the parts more accurately over the vehicle.

In terms of individual parts, the same trend occurs, namely there is a slight decrease in performance when ACMs are used, with the exception of the wheels where the circular Hough transform combined with ACM results in an increased coverage of the part, as shown in Table 6. To further strengthen the evaluation, the table also includes a statistical evaluation measure, the average F-score obtained for each part of interest for the ACM.

4.2. Comparison with the literature

Fig. 13 shows a comparison of the proposed approach with various approaches existing in the literature [9,13–16]. To our knowledge there are currently no solutions for vehicle parts

Table 5 Accuracy of bounding box (BB) and ACMs localization for rear view.

Vehicle Part	Percentage of area (%) in BB/ACM			Parts found (%) (BB/ACM)
	Sedan	SUV	Wagon	
Left wheel	93.8/100	93.6/100	92.8/100	100/100
Right wheel	95.6/100	93.6/97.3	95.6/100	100/100
Window	81/82.8	87.2/93	80.1/74.1	100/97.5
Logo	96.6/78.2	100/81	98.8/79.4	100/92.3
Rear light left	92.5/97.2	93.5/98.3	92.8/90.8	100/98.8
Rear light right	92.5/97.7	93.6/94.3	92.5/92.8	100/98.8
Average	92/92.6	93.6/93.9	92.1/90.71	100/97.9



Fig. 12. Bounding boxes and corresponding ACMs for different vehicle categories viewed from different sides.

localization that are adapted to the use of multiple viewpoints. Moreover, none of the above methods can be easily expanded to cope with multiple viewpoints. This is due on one hand to the specific implementation mechanisms used for the detection of parts. The parts are identified based on the position of round wheels [12,15]. These are not visible in front and rear views therefore making impossible to use the methods on such views. On the other hand, other solutions exploit the specificity of the data available.

For example, video data is employed to detect the entry and exit of a car in the scene in [14]. In the case of the current work, only static data is available. In [13], key points are detected based on the availability of camera parameters. The latter remain unknown, and unnecessary, for the solution introduced in this work. Other solutions are dedicated to one single part of interest (e.g. logo in [9], rear lamps in [16]). These factors make impossible the implementation of a method based on the literature that is capable to deal with

**Table 6**  
Average accuracy of localization for all views.

Vehicle part	Percentage of area (%) in BB	Percentage of area (%) in ACM	Parts found (%) BB	Parts found (%) ACM	F-Score for ACM
Wheel	92.6	98.9	100	100	1
Wind-shield	93.4	80.2	100	98.8	0.98
Window	94.2	94.1	100	100	1
Bumper	100	96.1	100	96	0.94
Handle	96.7	90.6	96.7	91.4	0.91
Head-light	94.7	76.8	100	92.7	0.93
Rear light	94.9	92.3	100	98.9	0.94
Grille	99.4	97.5	100	100	1
Mirror	97.9	74.3	100	86.7	0.81
Logo	99.2	90.8	100	93.3	0.89
Gas trap	99.6	95.4	100	94.3	0.86

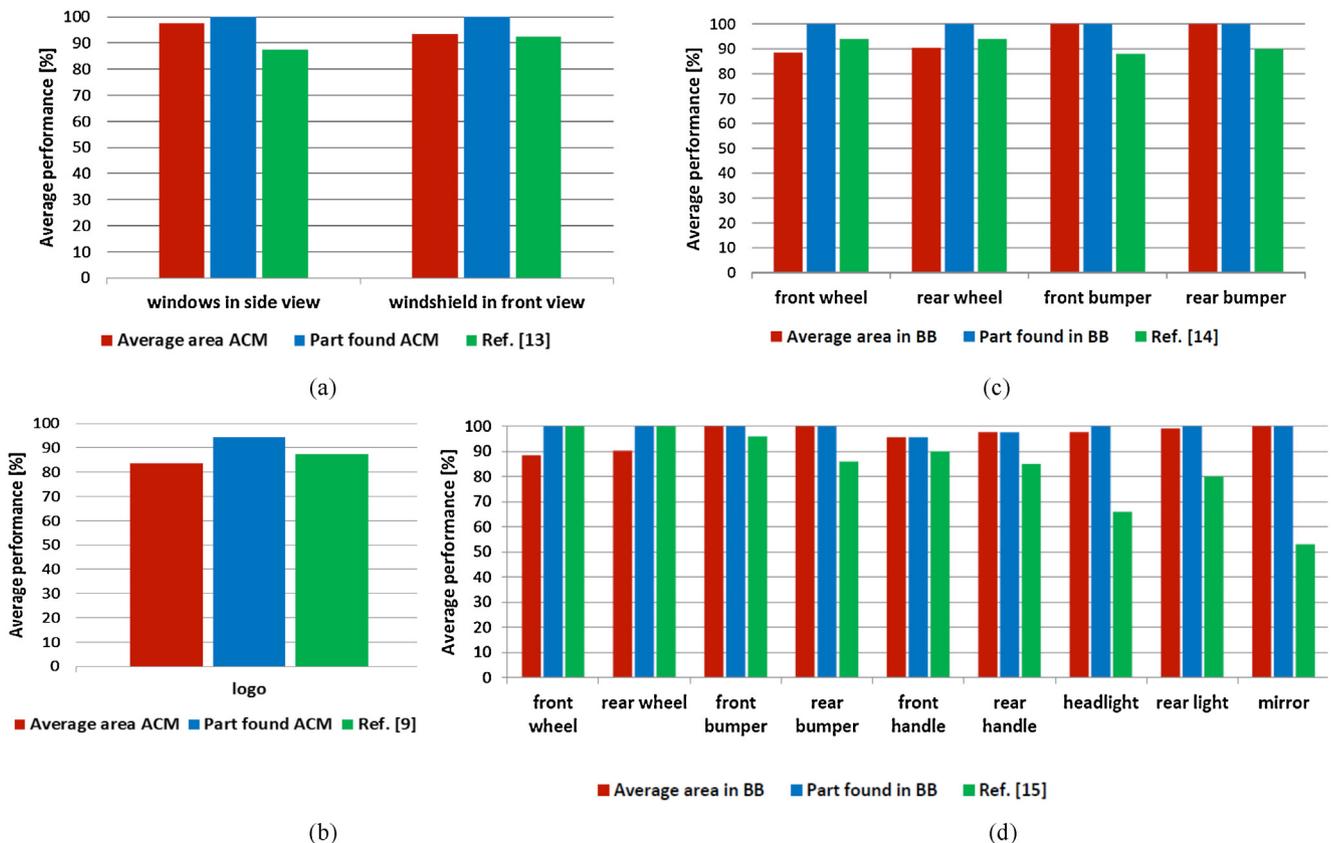
multiple viewpoints and all the parts of interest for direct comparison. Given these constraints, the results are only shown for the view and the parts that are reported in the respective work in the literature. For the methods that are bounding-box style, such as [14,15], a comparison is performed with the proposed bounding-box (BB) solution, while for those that report exact contours of parts [9,13], the comparison is performed with respect to the proposed ACM refinement that results in accurate contours.

The results reported in Fig. 13a are calculated based on the reported error rate in [13] for the defined point of view, and refer to the real contour of the windows and windshield. Because of this fact, they are compared with the proposed ACM solution. Other results reported in [13] are not directly comparable, because the performance is reported on regions containing multiple parts, such as grille, headlight and bumper region or the door panels, fender and wheel region. Similarly, the results in [9] in Fig. 13b, are reported for the exact contour of the vehicle logo and compared with the proposed ACM solution. The approaches in [14] and [15],

shown in Fig. 13c and d, respectively, identify parts as bounding boxes and are therefore compared with the bounding-box solution proposed the context of this work.

It can be seen that the proposed method, shown with red (percentage of area of the part within the ACM or BB) and blue (part found in ACM or BB), achieves better detection and localization rates, especially for parts that are visually more challenging to detect such as lights, door handles and mirrors than the solutions in the literature, shown in green. Furthermore, the same conclusion can be derived by comparing the F-score reported in [16] for the rear lamp. The reported value is 0.81, while the one obtained by the proposed solution 0.94, as shown in Table 6.

In terms of computation time, it takes for the proposed solution on average 0.01 s from the moment of reading an input image to the moment that the bounding boxes are displayed using a Matlab platform running on a Pentium Intel Core 2 duo at 3.0 GHz and 2.0 GB RAM. The computation of average models, including the computation of local minima and maxima and the adjustments, takes



**Fig. 13.** Comparison of performance with related work: (a) [13], (b) [9], (c) [14] and (d) [15].

**Table 7**  
Average computation time (ms).

	Bounding box approach [15]	Bounding box proposed solution
Avg. time (ms) per part	55–170	19

on average about 0.36 s per vehicle, but is performed offline. The fine tuning using ACMs comes at additional computation cost of about 4 s per vehicle or about 430 ms per part, which is not significant considering the amount of time required to displace robotic equipment. Table 7 compares the computation times per part with the solution reported in [15].

For the bounding-box style of localization, the proposed solution is faster than the similar style bounding box method proposed in [15], running on the same computer.

## 5. Conclusion

The paper presents a novel, improved bounding-box approach for the identification of the location of different vehicle parts from multiple views. It is based on a model of human visual attention and capitalizes on the correspondence between the location of parts and the projections on the axes of the obtained model. The bounding boxes are automatically adjusted for different categories of vehicles, for different views of each vehicle and for different exemplars within a category of vehicles. They were shown to fit better with the various vehicle parts than other solutions proposed in the literature. Moreover, the proposed method offers a very efficient and robust initialization stage to finely tune the bounding boxes to more accurate contours with ACMs and therefore to a more complete and accurate description of vehicle parts which none of the current solutions in the literature achieves.

As future work, the framework will be expanded for additional vehicle types and further improvements will be brought to cope with specific characteristics of different categories of vehicles, such as the different shapes and placement of the gas trap, or the different positions of door handles (e.g. towards the left side of the rear door for given wagon models). Other solutions for capturing the contour of parts will be studied as well to improve the performance of the proposed system.

## Acknowledgements

The authors acknowledge the support to this research from the Natural Sciences and Engineering Research Council of Canada under the Strategic Project Grant No. STPGP 381229-09, as well as the collaboration of Scintrex Trace Corp.

## References

- [1] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (11) (1998) 1254–1259, <http://dx.doi.org/10.1109/34.730558>.
- [2] T.F. Chan, L.A. Vese, Active contours without edges, *IEEE Trans. Image Process.* 10 (2) (2001) 266–277, <http://dx.doi.org/10.1109/ICCP.2012.6356188>.
- [3] Z. Sun, G. Bebis, R. Miller, On-road vehicle detection: a review, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (5) (2006) 694–711, <http://dx.doi.org/10.1109/TPAMI.2006.104>.
- [4] C.N.E. Anagnostopoulos, I.E. Anagnostopoulos, V. Loumos, E. Kayafas, A license plate recognition algorithm for intelligent transportation system applications, *IEEE Trans. Intell. Transp. Syst.* 7 (3) (2006) 377–392, <http://dx.doi.org/10.1109/TITS.2006.880641>.
- [5] M. Chacon, A. Zimmerman, License plate location based on a dynamic PCNN scheme, *Proc. Int. Jt. Conf. Neural Netw.* 2 (2003) 1195–1200, <http://dx.doi.org/10.1109/IJCNN.2003.1223862>.
- [6] L.M. Guo, X. Wang, Yuan, Car plate localization using pulse coupled neural network in complicated environment, in: Q. Yang, G. Webb (Eds.), *PRICAI 2006, Lecture Notes in Artificial Intelligence* 4099, 2006, pp. 1206–1210, [http://dx.doi.org/10.1007/11816171\\_139](http://dx.doi.org/10.1007/11816171_139).
- [7] K.-B. Kim, S.-W. Jang, C.-K. Kim, Recognition of car license plate by using dynamical thresholding method and enhanced neural networks, in: N. Petkov, M.A. Westenberg (Eds.), *CAIP 2003, Lecture Notes in Computer Science* 2756, 2003, pp. 309–319, [http://dx.doi.org/10.1007/978-3-540-45179-2\\_39](http://dx.doi.org/10.1007/978-3-540-45179-2_39).
- [8] H.J. Lee, Neural network approach to identify the model of vehicles, in: J. Wang (Ed.), *ISNN 2006, Lecture Notes in Computer Science* 3973, 2006, pp. 66–72, [http://dx.doi.org/10.1007/11760191\\_10](http://dx.doi.org/10.1007/11760191_10).
- [9] A.P. Psyllos, C.-N.E. Anagnostopoulos, E. Kayafas, Vehicle logo recognition using a SIFT-based enhanced matching scheme, *IEEE Trans. Intell. Transp. Syst.* 11 (2) (2010) 322–328, <http://dx.doi.org/10.1109/TITS.2010.2042714>.
- [10] C.-N.E. Anagnostopoulos, I.E. Anagnostopoulos, I.D. Psoroulas, V. Loumos, E. Kayafas, License plate recognition from still images and video sequences: a survey, *IEEE Trans. Intell. Transp. Syst.* 9 (3) (2008) 377–391, <http://dx.doi.org/10.1109/TITS.2008.922938>.
- [11] F. Tan, L. Li, B. Cai, D. Zhang, Shape template based side-view car detection algorithm, *Proc. IEEE Int. Worksh. Intell. Syst. Appl. (China)* (2011) 1–4, <http://dx.doi.org/10.1109/ISA.2011.5873380>.
- [12] R. Brehar, S. Nedevschi, L. Daian, Pillars detection for side viewed vehicles, *Proc. IEEE Conf. Intell. Comput. Commun. Process.* (2010) 247–250, <http://dx.doi.org/10.1109/ICCP.2010.5606430>.
- [13] W.W.L. Lam, C.C.C. Pang, N.H.C. Yung, Vehicle-component identification based on multiscale textural couriers, *IEEE Trans. Intell. Transp. Syst.* 8 (4) (2007) 681–694, <http://dx.doi.org/10.1109/TITS.2007.908144>.
- [14] W.-C. Chang, C.W. Cho, Real-time side vehicle tracking using parts-based boosting, in: *Proc. IEEE Conf. Syst. Man Cybernet. (Singapore)*, 2008, pp. 3370–3375, <http://dx.doi.org/10.1109/ICSMC.2008.4811818>.
- [15] A. Chávez-Aragón, R. Laganière, P. Payeur, Vision-based detection and labelling of multiple vehicle parts, *IEEE Trans. Intell. Transp. Syst.* (Washington, DC) (2011) 1273–1278, <http://dx.doi.org/10.1109/ITSC.2011.6083072>.
- [16] B. Tian, Y. Li, B. Li, D. Wen, Rear-view vehicle detection and tracking by combining multiple parts for complex urban surveillance, *IEEE Trans. Intell. Transp. Syst.* 15 (2) (2014) 597–606, <http://dx.doi.org/10.1109/TITS.2013.2283302>.
- [17] B. Tian, Y. Li, B. Li, F. Zhu, G. Xiong, An electronic police system with multiple vehicle parts model, *Proc. IEEE Conf. Serv. Operat. Logist. Inform. (Dongguan, China)* (2013) 281–286, <http://dx.doi.org/10.1109/SOLI.2013.6611426>.
- [18] A.-M. Cretu, P. Payeur, R. Laganière, Salient features based on visual attention for multi-view vehicle classification, *Proc. IEEE Comput. Intell. Meas. Syst. Appl. (Ottawa, Canada)* (2011) 1–6, <http://dx.doi.org/10.1109/CIMSA.2011.6059933>.
- [19] A.-M. Cretu, P. Payeur, Biologically inspired visual attention features for a vehicle classification task, *Int. J. Smart Sens. Intell. Syst.* 4 (3) (2011) 402–423.
- [20] Izmostock – Car Stock Photos, <http://www.izmostock.com/> (accessed January 2013).