HAL
open science

A visual attention focusing system using an active stereoscopic vision sensor

Yann Ducrocq, Shahram Bahrami, Luc Duvieubourg, François Cabestaing

To cite this version:

Yann Ducrocq, Shahram Bahrami, Luc Duvieubourg, François Cabestaing. A visual attention focusing system using an active stereoscopic vision sensor. International Conference on Image Processing Theory, Tools and Applications, IPTA'2010, Jul 2010, Paris, France. pp.id:139. hal-00521099

HAL Id: hal-00521099

https://hal.science/hal-00521099v1

Submitted on 25 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A visual attention focusing system using an active stereoscopic vision sensor

Yann Ducrocq [1], Shahram Bahrami [1], Luc Duvieubourg[2] and François Cabestaing[2]
[1] École d'Ingénieurs du Pas-de-Calais, Département Automatique, Campus de la Malassise, 62967 Longuenesse Cedex - FRANCE,
e-mail: yann.ducrocq@eipc.fr, shahram.bahrami@eipc.fr
[2] Laboratoire LAGIS - UMR CNRS 8146, Université des Sciences et Technologies de Lille, Cité Scientifique
Bâtiment P2, 59655 Villeneuve d'Ascq - FRANCE
e-mail: luc.duvieubourg@univ-lille1.fr, francois.cabestaing@univ-lille1.fr

***Abstract*— Humans have the capacity to move their eyes. Thanks to this capacity, they can orient their gaze to look at a relevant object inside a complex scene. In this paper, we have implemented a driver assistance application which tries to mimic this human capacity. We focus specially on highway driving situations, where the detection of obstacles must be done far away in front of the car. The implementation of the gaze control and orientation is obtained by an active vision system. We know that the human gaze is related to the visual attention which is a result of human perception and cognitive phenomenon. Several studies have shown that human perception and more specially the visual perception can be decomposed into a bottom-up process and a top-down process. Most of researches focused on the bottom-up process. In this work, in order to mimic the human behavior or at least improve vision systems, we use a new active stereovision setup and a model of the human visual perception based on the two previous approaches. Moreover, in a bottom-up approach, we add the depth information obtained with the stereoscopic sensor to the classical features used by other works. The top-down process is computed by the global knowledge of the scene and its features. Some results obtained by mean of a virtual road sequence, show the orientation of the field of view of the stereoscopic sensor toward relevant objects, given our criteria.**

***Keywords*— Stereoscopic catadioptric sensor, orientable field of view, active vision, visual attention, saliency**

I. Introduction

In the context of Advanced Driver Assistance System (ADAS), a lot of studies have been done during the last decade [1], [2], [3], [4]. Some of these studies have been devoted to environment sensors, *e.g.* radar, lidar, sonar or video-based sensors in order to provide information to the driver about the scene surrounding the car. The video-based sensors are limited in distance. The camera and its lens are not able to satisfy simultaneously a good accuracy of detection and a wide field of view to see the entire scene. If we use a large angle of view then we have a lack of accuracy in the detection process and consequently in the identification process. If we have a smaller field of view, the detection and identification processes will be better but we can miss obstacles in front of the car especially in curved road. From this analysis we have developed a new orientable stereoscopic sensor with a small field of view. This orientation possibility allows to scan the entire scene in front of the car. Specifically, we can orient the sensor toward the area of interest. Consequently, it is necessary to develop a strategy of visual perception for our active vision system in order to determine: **what to look** and **when to look**. From human visual behavior and specific criteria dedicated to our application, we develop an hybrid approach from the bottom-up and the top-down processes. The proposed approach selects the point of fixation from potential areas of interest by means of an evaluation function. It allows to obtain required information, dependent on our application goals, by controlling the active vision system.

The paper is organized as follows: Section 2 provides background information and literature review about visual attention. Section 3 deals with our hybrid approach. Our stereoscopic active vision sensor is described in section 4. Section 5 presents some results and evaluations from a sequence of synthetic images of road traffic. We conclude with a discussion about ways of improvement of the proposed approach.

II. Related work on field of view control

In order to set up a tool for driver assistance, the perception module has to give priority to relevant objects in the scene. The benefits that can be drawn are:

- A lower amount of data to analyze since the regions with relevant objects are limited,
- Better and faster identification of the objects in cluttered scene.

A. Field of view control based on risk assessment

In the context of assistance system for intelligent vehicle, Lattner proposed a knowledge-based risk assessment in order to identify objects which might be dangerous for the vehicle or other traffic participants [5]. He proposed a qualitative representation of the scene based on classes which describe the objects, the spatial and temporal relations between them and the road and traffic situation. An abstract description of the situation gives a risk pattern before a pattern matching, which notifies the risk assessment of the objects, and selects the region which should be focused. This system seems efficient, however it does not consider the characteristics of the human cognitive and visual system.

B. The human visual perception modeling

The human perception system and the related cognitive mechanisms have been studied both in synthetic faces with the

goal of improving the naturalness of animated virtual agents, and in road scene analysis to focus on relevant objects [6], [7].

1) The visual attention: For living beings, the relevant information is selected in the perceptual available data stream, before being prioritized processed by other stages. This process is called the selective visual attention. One aspect that emerges from studies in various fields like neuro-sciences and cognitive psychology is that the attention is controlled by the salience. The attention has a filter function that enables a gaze control on salient features of the scene, avoiding an exhaustive or random exploration when looking for information. For the development of our analysis scene tool, we can use this idea of visual attention to orient the field of view of the sensor to salient regions.

a) The saliency: an object is salient when it draws our attention. The salience depends on the observer, the objects and the context. An interesting study conducted by Landragin proposes a classification of the physical and cognitive factors, that make an entity salient in a linguistic utterance or in a visual scene [8]. Two main processes involved in the saliency mechanism are resumed below:

- *The bottom-up process* (or ascendant process) is driven by scene features, generally fast, spontaneous and automatic. The features cover information specific to the nature of an item, a singularity of an item in a set, and factors related to the cohesion and the structure of the scene [8]. Various stimuli contribute in putting forward an item when it is different from its neighbors on a given property.
- *The top-down process* (or descendant process) : some high-level controlled processes determine the visual attention because objects become salient according to a specific goal that we have to achieve or a succession of reasonings we can have. According to Landragin, cognitive factors in visual perception include the intention (the influence of the task ongoing on the perception), the attention (this is an attitude of preparation to the perception), the short-term and long-term memories [8], [9].

However, the human visual attention is guided by the two above processes. That leads to react to unpredicted event and to focus on the goals which were previously defined simultaneously. There exists a time-limited retroaction of the bottom-up process on the top-down process.

b) Personal factors affecting the visual attention: some factors related to the affect (stress, fear) influence the visual attention according to the personality of people. The effects of these factors are difficult to predict and are almost inordinate.

c) Inhibition process: there are some phenomenon of inhibition which control the saliency and therefore the visual attention. The attention is first shifted to the most salient location, then the saliency of this region decreases and the attention is shifted to the next most salient location. This inhibition has several dimensions: temporal, spatial and elimination of stimulus that appear not quite relevant (signal to noise rate).

2) Computational model of visual attention: The first computational models worked with static frames [10], [11] before to be extended to the analysis of dynamic scenes [12], [13], [14]. Meanwhile, descendant models (**relevant** map) have been added to the ascendant models (**saliency** map) to achieve so-called hybrid models [15], [16].

A computational model of bottom-up process, incorporating the biologically inspired ideas of Koch and Ullman, has been proposed by Niebur [10]. Thereafter Itti refined and improved this model [11]. The saliency is determined by how different a stimulus is from its surround, in many submodalities and at many scales. His method gives a saliency map in a several stages calculation.

Rosenholtz proposed a statistical approach. First a search display is represented in an appropriate feature space, then the saliency is computed by essentially performing a test for outliers [17]. This saliency model has been implemented to extract motion saliency from video and to predict where people look at through the windshield of a car.

Frintrop presented an application of visual attention mechanisms for the selection of point of interest in an arbitrary system. She used a 3D laser scanner mounted on a mobile robot. Her saliency-based region search indicates areas of interest of the scene. However, she didn't use this information to drive the robot to the interesting area or to identify the objects present in that region [18].

Michalke presents an attention system as a part of a biologically motivated advanced driver assistance system. He focuses on robustness enhancements in order to cope with the challenges a system is faced with when using saliency on real outdoor scenes [19].

In top-down attentional modulation, the knowledge could be modeled by a fuzzy logic algorithm [20], a bayesian network [16] or belief functions [21]. These models give additional information about the objects depending on the situation and the task-relevance of an entity [15]. This knowledge is used to determine the actions to make on the saliency map creation process. The information raise to the ascendant part of the system and the actions are asked by the descendant part. This process, often associated with a memorization process, runs on until a condition is verified, like the recognition of a situation.

III. The proposed model

The proposed model of our visual attention focusing system using an active stereoscopic vision sensor is displayed in figure 1.

From the images of the stereoscope we compute a saliency map. Also with the sequence of images, in an other process we detect the road and the obstacles to construct a map of the environment. The bottom-up process exhibits regions that are salient because of specific conspicuities, the top-down process emphasizing the features of the chosen scenarii. Both the saliency map and the observation strategy compete in an

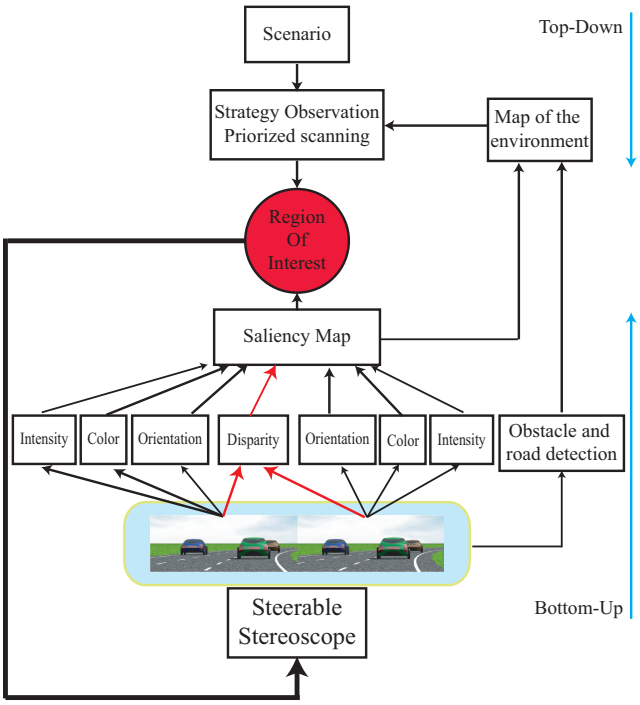


Fig. 1. The active visual system model

algorithm to search for the regions of interest (ROI). The coordinates (center of gravity) of these ROI are converted in a sequence of angles to control the orientation of the field of view of the stereoscope.

A. The bottom-up process

1) 2D information: We compute a saliency map to model the bottom-up process. This model uses stimuli from the scene: luminosity, color, orientation. For every characteristic, the visual system computes the differences between stimuli in order to see the most relevant ones. In our study, we use:

- the saliency based on the contrast, the luminosity and the color,
- the movement for the dynamic aspect,
- the orientation of the object which changes in time and appears more salient for human.

With each studied feature, we compute a conspicuity map (see [13]) for the right and the left stereoscopic images. This conspicuity map represents the information which differs from the others for a particular feature. For n features, we obtain n conspicuity maps and for a stereoscopic pair of images, we obtain $2n$ conspicuity maps.

2) 3D information: In addition to the previous 2D features used to make the saliency map, we add the depth information which is given by our stereoscopic vision system. This depth information is computed with a similarity-based adaptive neighborhood (SBAN) dense stereovision algorithm [22]. Figure 2 represents a pair of stereoscopic images (a) and the corresponding dense depth map (b). In this map, each pixel represents a distance information : light gray pixels belong to the nearest objects and dark gray pixels belong to the most distant objects.

a b

Fig. 2. Right and left stereoscopic images (a) and depth map (b)

In human vision, depth information plays an important role. Particularly, a driver always needs to use the distance information to locate his car to the others. In computer vision, in robotics for obstacle avoidance for example, it is important to know the distance between the sensor and the scene objects. Authors like Ouerhani [23] show that we can use the depth information to compute the saliency map. As it was made with the other features, we compute a conspicuity map based on the distance information. So, the objects closer than the others will be brought to light.

3) Combining 2D and 3D related conspicuity maps: To the $2n$ conspicuity maps computed from the color image for each feature, we add one conspicuity map based on the depth map. So, in order to compute the saliency map S, $2n+1$ conspicuity maps Ci (where i is a feature) have to be combined :

$$S = \sum_{i=1}^{2n+1} \gamma_i N(Ci) \tag{1}$$

Here N is a normalization operator [12] and γ is a factor used to adjust the weight of each parameter. In our tests, we give more importance to the movement and distance of objects than to the brightness and color. Their coefficient is therefore higher.

B. The task guided model

We can note many reasons why the analysis of the salience is not sufficient to control the orientation of the field of view :

- If we have some *a priori* information, it is better to integrate them into the model.
- The algorithm for computing the salience does not necessarily give the most interesting region to observe.
- Since the salience is also related to the will of human and the job they are doing, we must control the orientation of the field of view by both the bottom-up and top-down processes.

A machine does not have the capacity to decide for all encountered situations. We propose some context-based control mechanisms of the field of view. Of course, these mechanisms will influence the search of the most interesting region to explore.

1) Periodic imposed orientation of the field of view: To avoid the system to focus on the same region during a too long period and ignore other parts, we impose a periodic return of the field of view to the center of the road, in front of the car.

Another option would be a periodic total scan of the scene in front of the car.

2) Priorized scanning: On the assumption that the field of view is focused on the most salient region, if the salient region given by the bottom-up process changes frequently, the motorized stage of the sensor will also turn permanently. We will not only consider the most salient region, but several regions with the highest saliencies and affect them by a cost coefficient which would guide the orientation. On figure 3, suppose the current orientation is towards i, and the candidate regions to explore are j and k, with k being more salient than j. In this situation, shall we go to k then come back to j according to the saliency analysis, or go to j and then continue to k? To make a decision, we will compute a cost coefficient λ_{ij} which is proportional to the time required to bring the motorized stage from the angular position i to the angular position j: $\lambda_{ij} = \frac{\alpha_{ij}}{\Delta\alpha}$, with $\Delta\alpha$ the maximum angle of rotation authorized for the prism. α_{ij} is calculated from the mechanical specifications of the stereocopic sensor [24]. A priority coefficient p_i is then affected to the region to explore. This coefficient is obtained after an identification stage of the object in the highlighted area. We will give a greater coefficient, for example, to a pedestrian crossing the road than a vehicle driving in the same direction as ours. Finally, we compute $L_{ij} = \lambda_{ij} * \frac{1}{p_i}$. Thus, we can evaluate and compare the total costs $L_{ij} + L_{jk}$ and $L_{ik} + L_{kj}$ to choose the sequence of exploration. This analysis of the map of the environment provides a priority score $\bar{P}^i(t)$ for objects to observe.

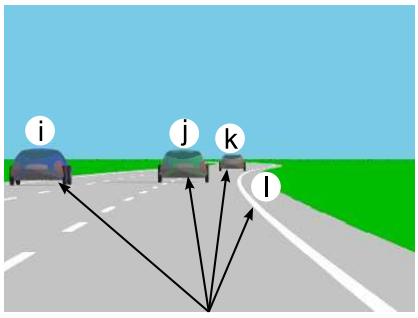


Fig. 3. Candidate directions to orient the field of view

C. Determining the region of interest to orient the stereoscope

The region of interest is the area of scene, in front of us, to which we orient the field of view of the sensor. This region is determined by an analysis of the saliency maps and the analysis of the environment map, built after the detection of the road and objects [25], [26].

We define four driving scenarii identified in the case of motorway driving :

- following : following a car in our lane ;
- overtaking : moving to the left side of the road, pass the car and going back to the right side of the road ;
- insertion : entering a highway based on the flow of traffic ;
- exit : exiting a highway.

According to these scenarii, we define sequences of actions and controls the driver has to do. The analysis of the road (GPS + road map + environment map), the intention of he driver (for example flashing) and the destination help us in this stage. The sequence of control is, per example, look at the left-look straight away-look at the right-look straight away in a passing situation. Obviously, this sequence is updated regularly, according the information arriving from the scene analysis.

The top-down aspect can influence the region of interest in two ways :

- guide the bottom-up analysis to highlight some objects (by adjusting the disparity map to reach more or less close objects) ;
- we modify the information search stage, in planning areas to observe, based on the will of the driver and the traffic situation.

We note $S^i(t)$ the saliency score of the subject i computed at frame t by the algorithm of saliency analysis. This saliency value is weighted by a function which decreases the score of the salience when the elapsed time from the observation of this object grows. We obtain $\sigma^i(t)$:

$$\sigma^i(t) = S^i(t)e^{-a(t-t_0^i)} \tag{2}$$

t_0^i is the time from the beginning of the observation of the object i. The weighting by the exponential function limit the importance of an object seen for some time. The coefficient a controls the decay of the exponential function. Then, we define also an average score of salience $\bar{\sigma}$, taking into account the historical values of salience of an object over several frames :

$$\bar{\sigma}^i(t) = \frac{1}{n} \sum_{j=0:n-1, n<t-t_0^i} \sigma^i(t-j) \tag{3}$$

The window size n and the coefficient a are adjusted to orient our sensor toward a new object that appears in the scene, even if the saliency value of the new object is lower than the saliency of the object already present for some time.

Finally, the region of interest arises from the comparison of the two scores :

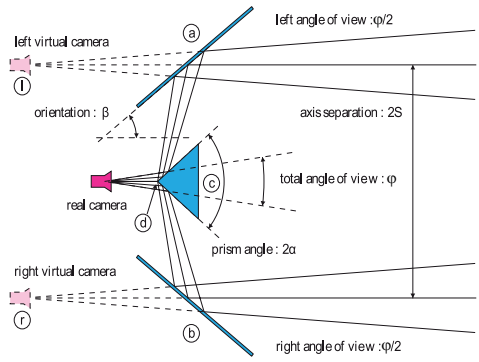
$$ROI = Max(\bar{S}^i, \bar{P}^i) \tag{4}$$

The position of the object with the largest score is analyzed to determine to what angle orient the sensor. This approach takes into account the predicted position of the object to be seen in terms of its dynamics, and mechanical limits of the sensor.

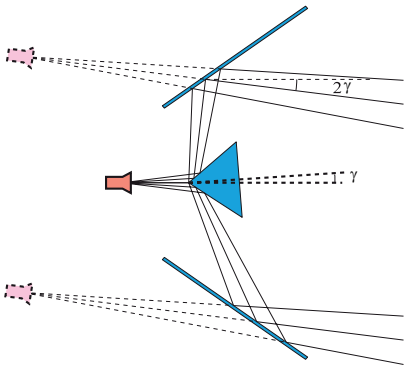
IV. Sensor description

A schematic top view of the stereoscopic sensor is presented in figure 4(a). The stereoscopic sensor is composed of a single camera, two lateral plane mirrors, labeled as (a) and (b)

figure 4(a), and a central prism with two planar reflective surfaces, labeled (c).



(a) without rotation



(b) rotation of angle γ

Fig. 4. Stereoscopic setup with and without rotation of the prism

We can find a description of the stereoscope and the characteristics of the active vision setup in [27], [28]. As this sensor has a small field of view, we can observe only a part of the scene in front of the vehicle. The sensor is designed to see two lanes of the road at one hundred meters. Without orientation of the prism and in straight road, the lane of our car and two half neighbored lanes in each side are visible. In order to observe the road scene and to identify the most salient objects, at long distance with a good accuracy, we orient the field of view of the system.

V. Experiments

For our tests, we need information about the scene. Particularly, we have to detect the road and the roadsides, evaluate the dynamic and the evolution of the detected objects. All these informations are expressed in x,y,z in our coordinate system reference, *id est* our car. This description takes into account the rotation of the field of view. Also, it is worth identifying the other objects in the scene in order to know or predict their evolutions. An obstacle or a pedestrian or a vehicle do not have the same behavior. These aspects are out of the scope of this paper. For our experiments, we use summary data from which this information is available.

1) Tested scenario: It is not possible for the moment to equip a car with our sensor laboratory prototype. So, in order to validate our approach, we generate video sequences with POVRAY [29]. Several scenari of road traffic were generated, with a screenplay by cars at different speeds on the motorway. In the frames shown by figure 5, a passing car situation, with three cars, is simulated. A car overtake, our car and another car driving ahead of us, before we overtake this last one.

For this sequence, we compute the saliency map as described in §III-A. We keep the three most salient elements. As we use summary data, we can identify each salient object. This information is used to evaluate the cost coefficient (see §III-B.2) of each salient region. According to this cost coefficient and the saliency score, we compute an attraction rate which guides the focus.

The result of our visual strategy is illustrated figure 5. From frames 1 to 4, we observe the most prior car which should be the car to monitor at that time. On frame 5, we orient the field of view straight away. In this configuration, we observe the lane where our car is located. After, we orient the field of view to the most prior car again. With our sensor, we can observe the entire scene at one hundred meters in three stages : a left orientation, a straight away orientation and a right orientation. When the sensor is oriented straight away, we observe the lane on which we are located. Every five frame, we start again and again the described procedure.

The developed method seems to be efficient in order to observe and detect correctly the salient objects of the scene. It also allows a periodic scan of the entire road. It is designed to be similar to a real situation of driving when an object appears in the scene while we follow an itinerary.

The proposed method is not yet completely implemented. But the first results on these generated video sequences with POVRAY seem to be encouraging. Note that the evaluation of our results on the focus of attention is very difficult because there is no ground truth.

VI. Conclusion and Future Work

This paper presents an approach of field of view control of a stereoscopic active vision system based on the human perception. We have presented the human cognitive perception and its two main mechanisms, the bottom-up and the top-down processes, studied in order to develop a strategy of observation. We tried to adopt a hybrid method by combining the two previous processes, in order to get a behavior of the system close to the human's one, or more concretely to improve our visual attention focusing system. The process of saliency map creation allows to consider the sensibilities of human eyes such as the color and the orientation. The top-down part influences the bottom-up part and modifies focus parameters got from the saliency map. Thus, the top-down approach guides the research of particular elements in the scene according to the situation. Finally, we have presented some first results with our developed system. In the future, it could be interesting to compare our results with an occulometer placed in a car driving simulator.

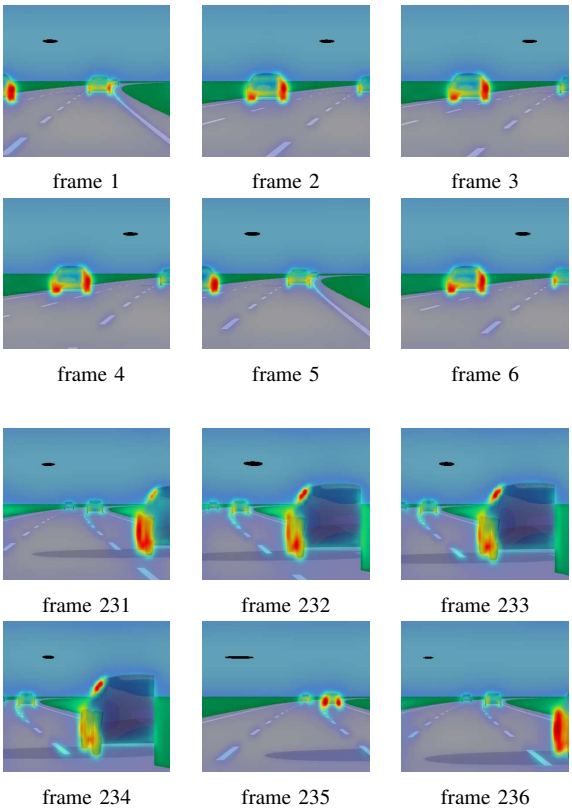


Fig. 5. Sequence of frames, simulation of a passing car situation. Frame 1-4 : monitoring the left car, frame 5 : switching straight away. frames 231-234 : monitoring a car appearing on our right, frame 235 : switching straight away

REFERENCES

[1] X. Clady, F. Collange, F. Jurie, P. Martinet : Object tracking with a pan tilt zoom camera : application to car driving assistance. In Proceedings of the IEEE International Conference on Robotics and Automation, ICRA'01, 2: 1653-1658, May 23-25th, Seoul, Korea, 2000.

[2] M. Pellkofer and E. Dickmanns : Ems-vision : gaze control in autonomous vehicles. In Proceedings of the IEEE Intelligent Vehicles Symposium, 296-301, 2000.

[3] M. Pellkofer, M. Lützeler, E. Dickmanns : Vertebrate-type perception and gaze control for road vehicles. In Springer Tracts in Advanced Robotics, 6: 271-288, 2003.

[4] S. Kammel and al.: Team AnnieWAY's autonomous system for the 2007 DARPA Urban Challenge. 25(9): Part II, 615–639, September, 2008.

[5] A.D. Lättner, I.J. Timm, M. Lorenz, O. Herzog : Knowledge-based risk assessment for intelligent vehicles. International Conference on Integration of Knowledge Intensive Multi-Agent Systems, 191-196, April 18-21, 2005.

[6] P. Mayenobe : (In French) Perception de l'environnement pour une gestion contextualisée de la coopération homme-machine. PhD Thesis, Université Blaise Pascal, France, December 2004.

[7] D. Herviou : (In French) La perception visuelle des entités autonomes en réalité virtuelle : application à la simulation de trafic routier, Thesis, Université de Bretagne Occiddentale, France, December 2006

[8] F. Landragin : Saillance physique et saillance cognitive, In Cognition, Représentation, Langage (CORELA), (revue électronique), 2004.

[9] R. C. Atkinson and R. M. Shiffrin : Human memory: A proposed system and its control processes. In Spence, K.W.; Spence, J.T., The Psychology of Learning and Motivation : Advances in Research and Theory, 2: 89-195, Academic Press, New York, 1968.

[10] E. Niebur and C. Koch : Control of Selective Visual Attention : Modeling the "Where" Pathway. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E.eds, Advances in Neural Information Processing Systems, 8: 802-808, MIT Press, Cambridge, MA, 1996.

[11] L. Itti, C. Koch, E. Niebur : A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11): 1254-1259, 1998.

[12] L. Itti : Models of Bottom-Up and Top-Down Visual Attention. PhD Thesis, California, 2000.

[13] J. Harel, C. Koch, P. Perona : Graph-Based Visual Saliency. NIPS2006-897, Advances in Neural Information Processing Systems, 19: 545-552, MIT Press, Cambridge, MA, 2007.

[14] J. Petit, R. Brémond, J.-P. Tarel : Une carte de saillance adaptée aux images HDR, Journées de l'Association Française d'Informatique Graphique (AFIG'08), 95-103, November 19-21, Toulouse, France, 2008.

[15] V. Navalpakkam and L. Itti : An Integrated Model of Top-Down and Bottom-Up Attention for Optimizing Detection Speed, Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2: 2049-2056, 2006.

[16] A. Torralba, A. Oliva, M.S. Castelhano, J.M. Henderson : Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. Psychological review, 113(4): 766-86. October, 2006.

[17] R. Rosenholtz : A simple saliency model predicts a number of motion popout phenomena, Vision Research, 39: 3157-3163, 1999.

[18] S. Frintrop, E. Rome, A. Nuchter, H. Surmann : Applying Attentional Mechanisms to Bi-modal 3D Laser Data. In International Workshop on Attention and Performance in Computer Vision (WAPCV '03), 25-30, Joanneum Research, Graz, Austria, 2003.

[19] T. Michalke, J. Fritsch, C. Goerick : Enhancing Robustness of a Saliency-Based Attention System for Driver Assistance. LNCS 5008: 43-55, 2008.

[20] Q. Andrieu, F. Dehais, A. Izaute, C. Lesire, C. Tessier : Towards a dynamic computational model of visual attention. International Conference on Humans Operating Unmanned Systems (HUMOUS'08), Brest, France, September, 2008.

[21] M. Veyret and E. Maisel : Simulation de la focalisation de l'attention visuelle : application à la simulation d'automobilistes virtuels, 17èmes Journées de l'AFIG, November, 2004.

[22] M. Pérez-Patricio, O. Colot, F. Cabestaing : A SBAN stereovision algorithm using hue as pixel similarity criterion. Computational Imaging and Vision, 32: 552-557, Springer, 2005.

[23] O. Ouerhani : Visual attention : from bio-inspired modeling to real time implementation. Université de Neuchâtel, Faculté des sciences, December, 2003.

[24] Y. Ducrocq, S. Bahrami, L. Duvieubourg, F. Cabestaing : An effective active vision system for gaze control. Proceedings of the 4th International Symposium on Visual Computing, ISVC 2008, LNCS 5359: 267-277, Las Vegas, NV, USA, Springer, December, 2008.

[25] R. Aufrère, F. Marmoiton, R. Chapuis, F. Collange, J.P. Deruin : Dtection de route et suivi de vhicules par vision pour l'ACC. Traitement du signal, 3(17): 233-248, 2000.

[26] D. Gruyer, V. Berge-Cherfaoui : Multi-objects association in perception of dynamical situation. Fifteenth Conference on Intelligent Robots and Systems, IROS'99, Kyongju, Korea, 17-21 October, 1999.

[27] L. Duvieubourg, S. Ambellouis, F. Cabestaing : Single-Camera Stereovision Setup with Orientable Optical Axes,Computational Imaging and Vision, 32: 173-178, Springer-Verlag, 2005.

[28] L. Duvieubourg, F. Cabestaing, S. Ambellouis, P. Bonnet : Long distance vision sensor for driver assistance, Proceedings of 6th IFAC Symposium on Intelligent Autonomous Vehicles, IAV'07, Toulouse, France, 2007.

[29] Persistence of vision raytracer. http ://www.povray.org/.