Universidad Carlos III de Madrid

e-Archivo
Institutional Repository

This is a postprint version of the following published document:

García, F.; Escalera, A.; Armingol, J. M. (2013). "Context aided multilevel pedestrian detection". *Information Fusion (FUSION), 2013 16th International Conference on*, Istanbul, 9-12 July 2013, pp. 2019-2024

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Context Aided Multilevel Pedestrian Detection

Fernando García, Arturo de la Escalera and José María Armingol
Intelligent Systems Lab.
Universidad Carlos III of Madrid
fegarcia@ing.uc3m.es

***Abstract*—The proposed work, depicts a novel algorithm able to provide multiple pedestrian detection, based on the use of classical sensors in modern automotive application and context information. The work takes advantage of the use of Joint Probabilities Data Association (JPDA) and context information to enhance the classic performance of the pedestrian detection algorithms. The combination of the different information sources with powerful tracking algorithms helps to overcome the difficulties of this processes, providing a trustable tool that improves performance of the single sensor detection algorithms.**

Context in a rich information source, able to improve the fusion process in all levels by the use of a priori knowledge of the application. In the present work multilevel fusion solution is provided for road safety application. Context is used in all the fusion levels, helping to improve the perception of the road environment and the relations among detections. By the fusion of all information sources, accurate and trustable detection is provided and complete situation assessment obtained, with estimation of the danger that involves any detection.

Keywords— Context, ADAS and Multilevel Application.

I. INTRODUCTION

The recent advances in computer vision and information technologies allow to create Advance Driver Assistant Systems (ADAS) to help to detect and warn the driver in advance in case of hazardous situations.

One of the main topics in the context of ADAS researches is the detection of pedestrians, whether by the use of modern and expensive sensors (e.g. radars), or inexpensive devices that introduce more uncertainties (e.g. computer vision). Among the classical approaches for pedestrian detection, the majority of works focus in the signal processing level and does not pay attention to the interactions among detections. But in real driving situations it is important to deal with the problems that arise in the data association process e.g: pedestrians merging into groups, as well as to take into account the interactions among them. It is in this field where this work tries to provide a step forward, by first providing fused pedestrian detection and later by studying the interactions of the pedestrians and providing efficient data association for the most challenging situations.

The approach tries to use the modern tools of information fusion to provide multilevel pedestrian detection based on multiple information sources.

II. STATE OF THE ART

Fusion approaches are common nowadays in Intelligent Transport Systems(ITS), due to the necessity of accurate and trustable detection. Among all a safety application, the pedestrian detection is one of the most important ones since it deals with the most vulnerable users of urban roads. In this context, computer vision and laser scanner are the two main sensors used for pedestrian detection. First offers a wide amount of unstructured information, this information is useful to provide classification, but its inferences lack of high reliability and have strong computational costs. On the other hand, information provided by the laser scanner is more trustable, but with limitations inherent of the technology (usually limited to one or four scan planes) and the excessive cost of modern devices with more planes. Thus the combination of both helps to overcome limitations inherent to each one.

As mentioned before, in ITS most of the works focus in the detection process i.e. levels 0 and 1:

In [1] and [2] authors present **centralized** works, combining medium level features form the different sensors (laser scanner and computer vision), different approaches are presented for final classification: Naïve Bayes, Gaussian Mixture Model Classifiers, Neural Networks, Fuzzy Logic Decision Algorithm, and Support Vector Machines(SVM).

Figure 1. IVVI 2.0 Researching platform for ADAS test used in the project

Decentralized schemes are frequent in ITS: [3] performs Adaboost based visual pedestrian detection and Gaussian Mixture Model (GMM) for laser scanner classification, a Bayesian decisor is used to combine detections at high level. In [4] multidimensional features for laser scanner are used, and the classical Histograms of Oriented Gradients (HOG) with SVM for computer vision detection, high level fusion is performed by Bayesian model. In [5] pattern matching approach is used for laser scanner, and stereovision with vertical projection of the human silhouette for computer vision detection, the fusion stage is based on a Global Nearest Neighbor (GNN) association algorithm.

Other approaches use the advantages of the different sensors in a different way. In [6] authors identify where pedestrians are more vulnerable, or difficult to be identified by the driver, using a laser scanner where the final pedestrian detection is done using the visual HOG features.

Context is a recent concept in information fusion, that takes advantage of the a priori knowledge of a given application to help in the fusion process. Some work already deals with the possibility of using context in vehicle applications: [7] use context in a Bayesian Network implementation, to determine the evolution of a detection probability of each track along time.[8] applies Bayesian methods for exploiting digital road-maps and realistic GMTI sensor modeling.

The work presented is a decentralized scheme, based on two independent low level classifiers (laser scanner and computer vision) and a final approach based on a powerful Multiple Target Tracking (MTT) algorithm, Joint Probabilities Data Association (JPDA). The decentralized approach represents robust configuration, able to provide detection even in extreme situations, where one of the two sensors is not available, and to provide robust detection by fusing the detections at high level. Furthermore, the JPDA [9] and [10] approach represents a highly adaptable algorithm, able to overcome difficult situations in the tracking stage. All the levels of the fusion process take advantage of the context information to enhance the fusion process.

III. GENERAL DESCRIPTION

Three sensors are available in the test platform IVVI 2.0m (Figure 1): Laser scanner, computer vision camera and a GPS with inertial measurements.

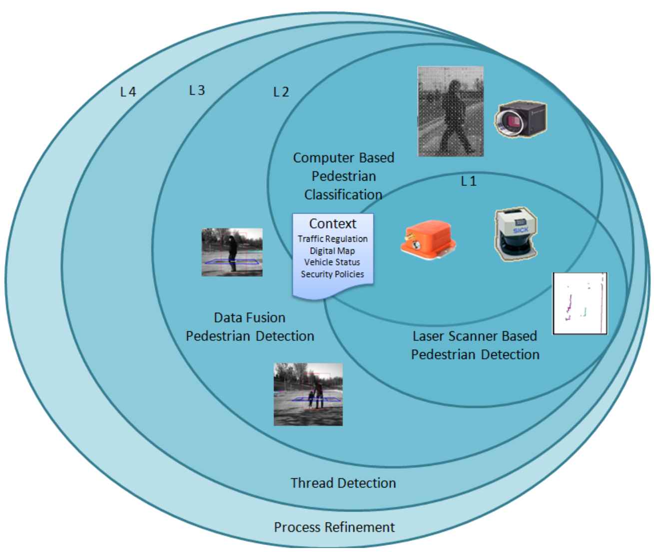

Figure 2. Application general digaram that represents the information included in each level.

The aim of the proposed work is to enhance the classical detections by providing a preliminary step in developing a fusion procedure for pedestrian detection. Two main factors were used to help in the fusion process. First, association process, based on JPDA, which proved to be very efficient in the association process facing difficult situations. Second, context is used to assist in the fusion process at all levels. At levels 0 and 1 it allows to increase the reliability of the detections. Lately at fusion levels 2 and 3 context is used to determine the threat that a given detection represents.

The work is a step forward in the purpose of developing a full architecture, able to provide fusion solution in all levels of the fusion process, taking into account all the available information, whether online (pedestrian detections, vehicle velocity, GPS position...) or offline (digital maps, safety regulations...). This way, the proposed method provides multilevel based information to enhance the security of the road users. Figure 2 depicts the general scheme of the proposal, all the levels are included, taking information from lower levels, and the sensors included on it.

In road safety applications, pedestrian detections are important, but the amount of pedestrians in urban environments is high. Thus it is important, to reduce the stress that the information of the pedestrians detections produces to the driver, by reporting only the safety threats i.e. pedestrians with high probability of interacting with the vehicle. This way the driver can focus in the driving process. These safety threats are estimated according to three aspects: danger estimation according to the distance to the car [5], location of the pedestrians (based on digital maps) and collision times.

The application provides solution in all fusion levels, taking advantage of the context information as follows:

Level 0. Related with data preprocessing, it deals with the alignment of the data and the synchronization. The online information is used to compensate the movement of the vehicle. First in the acquisition phase of the laser scanner points, and later in the extrapolation of the points to the moment where the images are acquired.

Level 1. Related with the object assessment, at this point the low level detections given by the laser scanner and the cameras are calculated. In both context information plays an important role. In the first it is used to detect the relevant obstacles, that fit the anthropological model of human beings. Later, it is also used to provide region of interest to the vision system, again based on anthropometric information.

Level 2. In this level final detections are provided according to the association process, based on JDPA Filter. Context information was used to adapt the filter to the specific situation of relevant detection. Important parameters, such as track creation or deletion policy, varies according to relevant context information, such as danger estimation, distance to the vehicle..

Level 3. At this level relevant information is used to provide threat detection, thus those pedestrian that represents a real danger to the vehicle are reported. Here context is mandatory, whether by offline information (e.g. traffic safety regulation and relevant distances) and online information (vehicle velocity, pedestrian location in digital maps, time to collision, etcetera)

Level 4. The process refinement on this application is also connected to the context information i.e. according to the information of threat provided in level 3, the tracking processes and the behavior of the system changes, allowing the system to focus in the relevant detections.

IV. Low Level Detection and Data Alignment

Low level detection and classification is based on laser scanner and vision. First is a more reliable sensor for obstacle detection, but it provides limited information, thus classification using the laser scanner is a difficult task. However, the reliability of it is used in the computer vision approach, to take advantage of the high reliability of the detections, to provide accurate obstacle detections, overcoming one of the difficulties of the computer vision approaches.

A. Laser Scanner Detection Algorithm

After the laser scanner retrieves the information, it is mandatory to correct the movement of the vehicle by using high accurate vehicle monitoring system. The platform used IVVI 2.0, has a high accuracy GPS augmented by inertial measurements [11]. This fusion methodology is used to retrieve online information of the movement of the vehicle. This information is used in the reconstruction of the points provided by the laser scanner (1) and the extrapolation of the laser scanner detection to the time of the image, thus providing time synchronization.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R(\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + T_v + T_0) \quad (1)$$

$$\text{where } T_v = \begin{bmatrix} vT_i \cdot cos(\Delta\varphi_i) \\ vT_i \cdot sin(\Delta\varphi_i) \\ 0 \end{bmatrix}, T_0 = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix},$$

R is the rotation matrix based on Euler angles, $\Delta\varphi_i$ corresponds to the increment in yaw angle at a given time, provided by the inertial sensor. Coordinates (x,y,z) and (x_0,y_0,z_0) are the Cartesian coordinates of a given detection of the laser scanner after and before to the movement compensation respectively. R is the rotation matrix, T_v the translation matrix according to the velocity of the vehicle, T_0 the translation matrix that represent the distance of the laser scanner to the inertial sensor. v is the velocity of the vehicle, T_i the time delay for a given point.

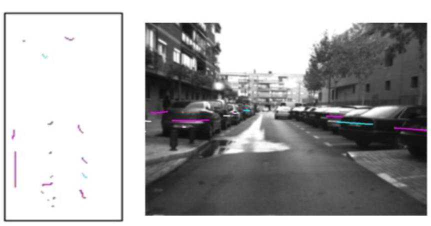

Figure 3. Polyline example, left only laser data, right data extrapolated to the camera.

Detection points are clustered according to the distance among them. After clustering, the obstacles are reconstructed using polylines [12] (Figure 3).

After movement compensation, pedestrian classification is performed. A deep study of the pattern of pedestrians was performed. A model was defined, based on the movements of the two legs while walking (Figure 4).

The pattern consists on consecutive polylines fulfilling several constraints regarding to angles and sizes [5]. Rotation of the pattern allows to extend the detection to lateral and diagonal movements.

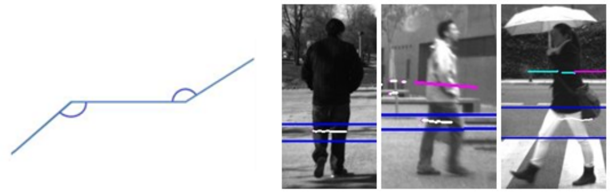

Figure 4: Laser scanner pattern (left), examples of leg movement (right).

Before pattern matching computation, candidates are selected according to the size, following anthropometric researches, that models human being using ellipses [13] and [14].

The lack of information (usually limited to one layer) is the biggest problem when using laser scanner classification. To overcome this problem a simple tracking stage was added based on Kalman Filter (KF). Final classification is based on the last 10 scans by a voting scheme. Consecutive scans are checked to eliminate false detections according to the behavior of the tracks. This misdetection filtering stage makes use of the contextual anthropometric information available (i.e. pedestrian velocities and sizes) to eliminate misdetections.

B. Computer Vision Detection Algorithm

Obstacles detected by the laser scanner are provided to the computer vision algorithm using pin-hole model, and with an accurate extrinsic parameter calibration, using the equation given in (1) adapted to the coordinate changes by the position of the two sensors (Figure 5).

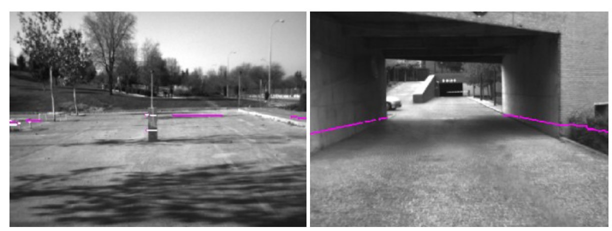

Figure 5. Laser scanner detections extrapolated to the camera with accurate precision.

Later, the obstacles with size according to human being are used to create regions of interest (Figure 6). This way the reliability of the laser scanner reduces the false positives of the image algorithms adding certainty to the detections.

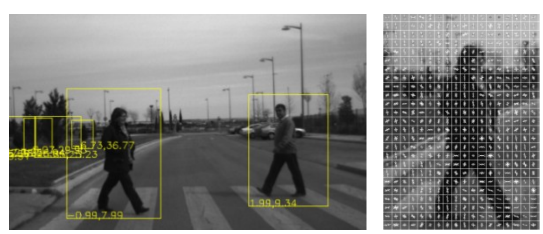

Figure 6. Examples of region of interest creation (left), with the distances to them highlighted, and HOG features calculation in a ROI(right).

Finally computer vision classification is performed by the HOG features approach and SVM [15].

V. TRACKING AND DATA ASSOCIATION

Data association and tracking was performed by a Joint Probabilistic Data Association Filter ([9] and [10]), This system was augmented by contextual information that allows the system to focus in those pedestrians that interacts with the vehicle.

JPDA denotes the association event ${\theta_k}_j^m$ that associates measurement m to track j as:

$$P(\theta|Z_k) = \frac{1}{c} p(z_k|\theta, X_k) P(\theta|X_k) \tag{2}$$

where c is the normalization constant, X_k is the target state vector. $P(\theta|, X_k)$ is the probability of the assignment θ conditioned to the sequence of the target state vector.

Assuming the M=2 dimensional Gaussian association likelihood for all the measurements to the target, the joint probability of a single measurement j to the target i is:

$$g_{i,j} = \frac{1}{(2\pi)^{M/2}\sqrt{|S_{ij}|}} e^{-\frac{{d_{i,j}}^2}{2}} \tag{3}$$

where $d_{i,j}$ is the distance between the prediction and the observation. The Cartesian approach was used, hence $\sqrt{|S_{ij}|} = \sigma_x\sigma_y$ as given per laser scanner obstacle detection test.

The resulting $P(\theta|Z_k)$ is:

$$P(\theta|Z_k) = {P_D}^{M-n} (1 - P_D)^n {P_{FA}}^{m_k-(1-M)} \prod_{j=1}^{m_k} g_{i,j} \tag{4}$$

where P_D is the detection probability, P_{FA} is the false alarm probability. n is the number of assignments to the clutter, m_k is the number of detections and M is the number of targets.

For the present approach, m_k is the number of observed pedestrians in a given scan time, with the information about the classification of both subsystems. Thus a detection Z_i it is obtained from the sensor defined as $[\, x_i , y_i , c_i , l_i \,]$ where (c_i, l_i) are the Boolean values that define the positive or not detection (TRUE or FALSE) by the computer vision approach or laser scanner respectively.

Finally all possible combinations that are included in the region of the gate (square gate), for a given track, are weighted using the likehood of the given association calculated using the joint probabilities:

$$R_k = \sum_{i=1}^{m} \left[\frac{P(\theta|Z_k)}{C} \left(Z_{i,k} - H_k \hat{X}_{k|k-1} \right) \right] \tag{5}$$

where R_k is the innovation covariance for the Kalman Filter of a given track at a given moment k, $\hat{X}$ is the state estimation of the KF, H is the state transition matrix of the KF, Z is the observation and C is the normalization factor that makes the sum of $P(\theta|Z_k)$ for all k in the gate equal to 1.

Although JPDA is a not a novel approach, its implementation represented a step forward in relation to other pedestrian detection approaches (e.g. [5]), typically based on other association algorithms i.e. Global Nearest Neighbors (GNN). Specifically, JPDA provide better solution in some specific situations:

- Double detections. It is when a laser scanner provides several regions of interest that include the same pedestrian, it can be caused by the perspective change or spurious measurements. Here only the most likely association has a weight considerable in the updating process to avoid any error in the tracking.
- Clustering errors. These usually happens when several pedestrians walk very close to each other, thus the laser scanner is unable to separate them. In this situation the association algorithm helps in the process of merging and separating, the detected observation helps in the updating process of all the tracks, allowing a more accurate tracking process. Similar situation happen due to close proximity of pedestrians.

A. *Track creation and deletion policy*

To perform track management, two types of tracks were defined: *consolidated* and *non-consolidated*. *Consolidated* refers to tracks with positive detection reported by both subsystems. *Non-consolidated* refers to tracks detected by a single subsystem, thus with not enough certainty to be reported. Although *non-consolidated* tracks are not trustable enough to be reported, it is important to keep the track of them, allowing them to evolve to *consolidated* tracks (e.g. when a pedestrian enters camera field of view).

A new track is created when a given measurement is out of any of the gates for the existing tracks.

A track is eliminated if a given track does not receive positive detection by any of the subsystems for a given number of the detections, this is denominated *maintenance*. The track logic defined specifies that a given measurement can be used only for the maintenance of a single track. Thus when a track falls in more than one gate, only the match with higher joint probability is used for maintenance. However, in the filter updating process, this observation is used in all tracks on which gates the observation falls.

The number of no-detections necessary for a given track depends on the threat priority of the track. As it is depicted in next section, the detections are labeled according to the danger that they low, medium and high. The number of no-detections change according the label of the track, so priority threat tracks have more latency. This number was empirically chosen to provide accurate tracking.

VI. PRIORITY TRACK ESTIMATION

Three aspects were taken into account for estimating the threat of a given track: danger estimation, distance to the road and time to collision.

A. Danger estimation

The danger estimation is done by evaluating the value of a danger function which depends on a pedestrian detection in relation to the movement of the vehicle [5]. Taking into account the average response time of drivers, the distance to completely stop the vehicle and condition of the road, the danger function of a detection is estimated to be:

$$f(r)=\begin{cases} e^{-\lambda(r-d_r)}, for \quad 80 \le r \le d_r \\ 1 \qquad\qquad , for \quad d_r < r \le 0 \end{cases} \tag{6}$$

where r is the distance of the car to the pedestrian and d_r is the reaction distance, defined as $d_r = vt_r$. Here v is the vehicle velocity and t_r is the response time that the driver needs to response to a given stimulus, which value was estimated according to anthropometric values in 0.66 seconds. λ is a parameter defined by:

$$\lambda = \frac{-\ln 0.6}{(d_b - d_r)} \tag{7}$$

where d_b is the braking distance, i.e. the distance that the vehicle covers since an alarm is triggered until it completely stops, defined as (8):

$$d_b = d_r + \frac{v^2}{\eta \mu g} \tag{8}$$

where η is the correction to the friction coefficient μ according to the state of the vehicle and the road, and g is the gravity force.

B. Distance to the road

Using digital maps, the distance of a given detection to the road where the vehicle is moving can be calculated. Using this distance it is possible to give an estimation of the danger that involves a given detection, based on an accurate location of the pedestrian. The accurate location is obtained from the laser scanner and the vehicle location by GPS with inertial measurement available in the IVVI 2.0 [11].

The estimation $(est(d))$ of the danger is modeled in a Gaussian manner, using width of the road as standard deviation:

$$est(d) = e^{-\frac{d^2}{2\sigma^2}} \tag{9}$$

where d is the distance to the center of the road and σ is width/2 of the road. By this estimation, pedestrians inside the road have danger estimation bigger than 0.6.

C. Collision estimation

The last parameter to take into account is the collision estimation with the pedestrian, given the velocities of both the vehicle and the pedestrian. According to [16], the collision solution can be obtained as follows:

The estimation of the state of the targets, defined by its position (x,y), and velocity vectors (vx,vy) is obtained by the KF. On the other hand, information from the GPS with inertial system allows estimation of the same kinematic information related to the vehicle. Thus the diagram shown in Figure 7 can be used to calculate the collision using (10)-(12) based on the aforementioned information.

$$x_c = \frac{(y_2 - y_1) - (x_2 \tan\theta_2 - x_1 \tan\theta_1)}{(\tan\theta_1 - \tan\theta_2)} \tag{10}$$

$$y_c = \frac{(x_2 - x_1) - (y_2 \cot\theta_2 - y_1 \cot\theta_1)}{(\cot\theta_1 - \cot\theta_2)} \tag{11}$$

where x_c and y_c are the coordinates of the collision point.

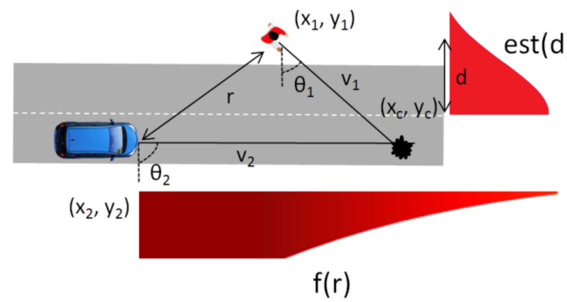

Figure 7. Representation of the trajectories, with the visual representation of the estimation of the danger for road proximity (est(d)) and the danger estimation according to the vehicle velocity (f(r)).

Once the collision point has been calculated, the next step is to calculate the time of each of the obstacles to this point, according to their respective velocity. When the time coincides this is considered the time to the collision. But in order to provide a safety margin, it is established a δ that defines the security margin, as depicted in equation (12). The higher the $\delta$the more conservative is the approach:

$$ttc\_d = |TX1 - TX2| < \delta \tag{12}$$

Again, a Gaussian manner estimation was created to determine the danger that involves a given detection, called collision estimation (ce), according to the time distance (ttc_d) and using the security margin δas the standard deviation:

$$ce(d) = \begin{cases} e^{-\frac{ttc\_d^2}{2\delta^2}}, for \quad collision \\ \qquad trajectories \\ \\ 0.4 \quad , for \quad no\ collision \\ \qquad trajectories \end{cases} \tag{13}$$

No collision trajectories are also taken into account with ce=0.4 to allow detection of possible danger pedestrians that are not colliding with the vehicle e.g: A pedestrian walking parallel to the vehicle within the road.

VII. THREAT PRIORITY

Threat priority classification has been finally explored as a decision-making solution using expert knowledge that has been presented in this work. The use of an expert module that includes knowledge, data and decision-making can be conducted by fuzzy logic. So, danger estimation, distance to the road and time to collision are easily combined for estimating the priority classification of a given track as low, medium or high threat. The decision-making allow a safe solution to determine when a given detection represent a threat to the vehicle or not, highlighting the individual knowledge of each danger estimation studied in this work.

VIII. TEST

Test were performed in three scenarios, test scenario, urban and interurban. First structured scenarios were used in which the detections are easy, thus low *false positive* rate is expected. These scenarios were useful to configure the detection subsystems. Next interurban scenarios are real scenarios, with low *false positive* rate expected due to the absence of other obstacles. Finally urban scenarios are less structured with high amount of obstacles that can lead to higher *false positive* rates. More than 10,000 frames where used with more than 4,000 pedestrians detections involved. Results are shown in Table 1.

	Camera		Laser Scanner		**Fusion**	
	% Pos	% Err	% Pos.	% Err	**% Pos**	**% Err**
Test	78.01	5.19	79.71	6.23	**82.42**	**0.89**
Interurban	73.19	3.91	70.35	16.96	**78.90**	**6.53**
Urban	67.72	6.72	73.61	16.72	**81.76**	**1.95**
Total	72.97	5.27	74.56	13.3	**82.29**	**1.11**

Table 1. Results in the different sequences. Pos for positive detections. Err for false positive errors in the detections.

Figure 8.Detection Examples with the estimation of the threat priority highlighted, green for detection low or medium and red for high.

IX. CONCLUSIONS

Results obtained in different scenarios, as well as the whole set of tests, are depicted in Table 1. Previously presented performances of the different subsystems are also included independently to allow to contrast the performance of the whole system and each system independently.

As depicted in Table 1, the results obtained increased and demonstrated the viability of the fusion and how it improves the overall performance of the system. The improvements of the system are summarized in the following points:

•The rate of positive detection increased in all tests. This increment is even better in the worst case scenarios for the laser scanner such as urban or inter-urban environments.

•It is also remarkable the improvements in the false positive rate. The results reach to 1%, providing the proof that the system fulfills the main requirements of a safety application: reliability. It is in this point where the improvement of the fusion process is most remarkable.

Finally Figure 8 depicts how thanks to the thread priority detection it is possible to reduce the load of the detections provided to the driver, since only those with high probability of interact with the vehicle are reported.

ACKNOWLEDGMENTS

This work was supported by the Spanish Government through the Cicyt projects (GRANT TRA2010-20225-C03-01) and (GRANT TRA 2011-29454-C03-02). CAM through SEGVAUTO-II (S2009/DPI-1509) .

REFERENCES

[1] C. Premebida, O. Ludwig, M. Silva, and U. Nunes, “A Cascade Classifier applied in Pedestrian Detection using Laser and Image-based Features,” *Transportation*, pp. 1153–1159, 2010.

[2] C. Premebida, O. Ludwig, and U. Nunes, “LIDAR and Vision-Based Pedestrian Detection System,” *Journal of Field Robotics*, vol. 26, no. Iv, pp. 696–711, 2009.

[3] C. Premebida, G. Monteiro, U. Nunes, and P. Peixoto, “A Lidar and Vision-based Approach for Pedestrian and Vehicle Detection and Tracking,” *IEEE Intelligent Transportation Systems Conference ITSC*, pp. 1044–1049, 2007.

[4] L. Spinello and R. Siegwart, “Human detection using multimodal and multidimensional features,” *2008 IEEE International Conference on Robotics and Automation*, pp. 3264–3269, 2008.

[5] F. Garcia, A. de la Escalera, J. M. Armingol, J. G. Herrero, and J. Llinas, “Fusion based safety application for pedestrian detection with danger estimation,” in *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, 2011, pp. 1–8.

[6] A. Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and H. G. Jung, “Localization and analysis of critical areas in urban scenarios,” in *Intelligent Vehicles Symposium, 2008 IEEE*, 2008, pp. 1074–1079.

[7] B. Jida, R. Lherbier, M. Wahl, and J. C. Noyer, *Bayesian Networks and Probabilistic Data Association Methods for Multi-Object Tracking Application to Road Safety*. 2008.

[8] M. Mertens and M. Ulmke, *Ground Moving Target Tracking with context information and a refined sensor model*, vol. 1, no. 5. 2008, pp. 1497–1504.

[9] Y. Bar-Shalom and X.-R. Li, *Estimation and tracking - Principles, techniques, and software*. Boston: Artech House.

[10] S. Blackman and R. Popoli, *Design and analysis of modern tracking systems(Book)*. 1999.

[11] E. D. Martí, D. Martín, J. García, A. de la Escalera, J. M. Molina, and J. M. Armingol, “Context-aided sensor fusion for enhanced urban navigation.,” *Sensors (Basel, Switzerland)*, vol. 12, no. 12, pp. 16802–37, Jan. 2012.

[12] F. García, F. Jiménez, J. E. Naranjo, J. G. Zato, F. Aparicio, J. M. Armingol, and A. de la Escalera, “Environment perception based on LIDAR sensors for real road applications,” *Robotica*, vol. FirstView, pp. 1–9, 2011.

[13] R. J. Skehill, M. Barry, and S. Mcgrath, “Mobility Modelling with Empirical Pedestrian and Vehicular Traffic Characteristics,” *WSEAS TRANSACTIONS on COMMUNICATIONS*, vol. 4, no. 10, 2005.

[14] *Highway Capacity Manual 2000*, no. 193. Transportation Research Board, National Academy of Sciences, 2000.

[15] N. Dalal and W. Triggs, “Histograms of Oriented Gradients for Human Detection,” *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05*, vol. 1, no. 3, pp. 886–893, 2004.

[16] F. Jiménez, J. Naranjo, and F. García, “An Improved Method to Calculate the Time-to-Collision of Two Vehicles,” *International Journal of Intelligent Transportation Systems Research*, vol. 11, no. 1, pp. 34–42, 2013.