

TECHNICAL UNIVERSITY OF CLUJ-NAPOCA, CLUJ-NAPOCA, ROMÂNIA  
FACULTY OF ELECTRONICS, TELECOMMUNICATIONS  
AND INFORMATION TECHNOLOGY  
and  
INSTITUT NATIONAL DES SCIENCES APPLIQUEES, ROUEN, FRANCE  
LABORATOIRE D'INFORMATIQUE, DE TRAITEMENT DE L'INFORMATION  
ET DES SYSTEMES

# Contributions to the Information Fusion. Application to Obstacle Recognition in Visible and Infrared Images

Ph.D. Student: **Anca DISCANT (épouse Apătean)**

Ph.D. Advisor: Professor **A. Benshair**  
Institut National des Sciences Appliquées, Rouen, France

Ph.D. Advisor: Associate Professor **A. Rogozan**  
Institut National des Sciences Appliquées, Rouen, France

Ph.D. Advisor: Professor **C. Rusu**  
Technical University of Cluj-Napoca, Romania



---

## ABSTRACT

The interest for the intelligent vehicle field has been increased during the last years, must probably due to an important number of road accidents. Many accidents could be avoided if a device attached to the vehicle would assist the driver with some warnings when dangerous situations are about to appear. In recent years, leading car developers have recorded significant efforts and support research works regarding the intelligent vehicle field where they propose solutions for the existing problems, especially in the vision domain. Road detection and following, pedestrian or vehicle detection, recognition and tracking, night vision, among others are examples of applications which have been developed and improved recently. Still, a lot of challenges and unsolved problems remain in the intelligent vehicle domain.

Our purpose in this thesis is to design an Obstacle Recognition system for improving the road security by directing the driver's attention towards situations which may become dangerous. Many systems still encounter problems at the detection step and since this task is still a work in progress in the frame of the LITIS laboratory (from INSA), our goal was to develop a system to continue and improve the detection task. We have focused solely on the fusion between the visible and infrared fields from the viewpoint of an Obstacle Recognition module. Our main purpose was to investigate if the combination of the visible-infrared information is efficient, especially if it is associated with an SVM (Support Vector Machine)-based classification.

The outdoor environment, the variety of obstacles appearance from the road scene (considering also the multitude of possible types of obstacles), the cluttered background and the fact that the system must cope with the moving vehicle constraints make the categorization of road obstacles a real challenge. In addition, there are some critical requirements that a driver assistance system should fulfil in order to be considered a possible solution to be implemented on board of a vehicle: the system cost should be low enough to allow to be incorporated in every series vehicle, the system has to be fast enough to detect and then recognize obstacles in real time, it has to be efficient (to detect all obstacles with very few false alarms) and robust (to be able to face different difficult environmental conditions).

To outline the system, we were looking for sensors which could provide enough information to detect obstacles (even those occluded) in any illumination or weather situation, to recognize them and to identify their position in the scene. In the intelligent vehicle domain there is no such a perfect sensor to handle all these concerned tasks, but there are systems employing one or many different sensors in order to perform obstacles detection, recognition or tracking or some combination of them. After comparing advantages and disadvantages between passive and active technologies, we chose the proper sensors for developing our Obstacle Detection and Recognition system. Due to possible interferences among active sensors, which could be critical for a large number of vehicles moving simultaneously in the same environment, we concentrate on using passive sensors, which are non-invasive, like cameras. Therefore, our proposed system employ visible spectrum and infrared spectrum cameras, which are relatively chosen to be complementary, because the system must work well even under difficult conditions, like poor illumination or bad-weather situations (such as dark, rain, fog).

The monomodal systems are adapted to a single modality, either visible or infrared and even if they provide good recognition rates on the test set, these results could be improved by the combined processing of the visible and infrared information, which means in the frame of a bimodal system. The bimodal systems could take different forms in function of the level at which the information is combined or fused. Thus, we propose three different fusion systems: at the levels of features or at the

level of SVM's kernels, or even higher, at the level of matching-scores provided by the SVM. Each one of these systems improves classification performances comparing to the monomodal systems. In order to ensure the adaptation of the system to the environmental conditions, within fusion schemes the kernels, the matching-scores and the features were weighted (with a sensor weighting coefficient) according to the relative importance of the modality sensors. This allowed for better classification performances. In the frame of the matching-scores fusion there is also the possibility to dynamically perform the adaptation of the weighting coefficient to the context.

In order to represent the obstacles' images which have to be recognized by the Obstacle Recognition system, some features have been preferred to encode this information. These features are obtained in the features extraction module and they are wavelet features, statistical features, the coefficients of some transforms, and others. Generally, the features extraction module is followed by a features selection one, in which the importance of these features is estimated and only the ones that are most relevant will be chosen to further represent the information. Different features selection methods are tested and compared in order to evaluate the pertinence of each feature (and of each family of features) in relation to our objective of obstacle classification. The pertinence of each vector constructed based on these features selection methods was first evaluated by a KNN ( $k$  Nearest Neighbours) (with the number of neighbours  $k = 1$ ) classifier, due to the simplicity in its usage: it does not require a parameter optimization process (as the SVM does).

To increase the accuracy of the classification, but also to obtain a powerful classifier, more parametrizable for the proposed fusion schemes, the KNN one was later (after the best features selection method have been chosen on the training set and the most relevant features have been selected) replaced by a SVM classifier. Because there is not known beforehand which combination of the SVM hyper-parameters is the most appropriate for a certain classification problem, an operation of model search, performed by 10 folds cross-validation, provides the optimized kernel for the SVM to be used on each fusion schemes and on each feature vector we considered.

Finally, we tested our features extraction, features selection and the proposed fusion schemes for a 4-class problem, thus discriminating between vehicles, pedestrians, cyclists and background obstacles. The results have proven that all bimodal visible-infrared systems are better than the monomodal ones, thus the fusion is efficient and robust since it allows for improving the recognition rates. In addition, features selection scheme provides smaller vector comprising only the most relevant features for the classification process. This reduction of the feature-vector dimension besides providing higher accuracy rates, allows the reduction of the computation time which is crucial in this type of application.

**Keywords:** Fusion, Infrared cameras, Features extraction, Features selection, Support Vector Machine, Kernels, Matching-scores, Hyper-parameter optimization, Model search, 10 folds cross-validation.

## REZUMAT

Interesul pentru domeniul vehiculelor inteligente a crescut simțitor în ultimii ani, cel mai probabil datorită numărului mare de accidente rutiere înregistrate. Majoritatea accidentelor ar putea fi evitate dacă autovehicolul ar avea atașat un dispozitiv de asistență a șoferului care să furnizeze semnale de atenționare atunci când intervin situații periculoase. Recent, marile companii de autovehicule au înregistrat eforturi semnificative și suport pentru cercetarea și soluționarea problemelor din acest domeniu al vehiculelor inteligente, propunând soluții pentru problemele existente, mai ales în domeniul viziunii computerizate. Detecția și urmărirea drumului, detecția pietonilor sau a vehiculelor, recunoșterea și urmărirea acestora, viziunea pe timp de noapte sunt câteva dintre aplicațiile dezvoltate și îmbunătățite recent. Cu toate acestea, încă există foarte multe provocări și probleme nerezolvate în domeniul vehiculelor inteligente.

Scopul nostru în această teză este de a descrie un sistem de recunoaștere a obstacolelor destinat îmbunătățirii securității rutiere prin direcționarea atenției șoferului înspre situațiile posibil periculoase. Multe sisteme încă întâmpină dificultăți în pasul de detecție și deoarece acest pas este încă în progres în cadrul laboratorului LITIS al INSA, scopul nostru a fost de a dezvolta un sistem care să continue și să îmbunătățească procesul de detecție. Ne-am concentrat doar pe partea de fuziune a câmpurilor vizibil și infraroșu din punct de vedere al unui modul de recunoaștere a obstacolelor. Obiectivul nostru principal a fost să investigăm dacă combinarea informației vizibil-infraroșu este eficientă, în special dacă este asociată cu o clasificare pe baza SVM (Support Vector Machine, în română - mașină cu suport vectorial).

Mediul exterior, multiplele posibilități de apariție a obstacolelor în scena rutieră (considerând de asemenea și multitudinea tipurilor de obstacole posibile), fondul foarte încărcat pe care pot să apară aceste obstacole și faptul că sistemul trebuie să considere și constrângerile de mișcare ale vehiculului determină categorizarea obstacolelor rutiere să fie o adevărată provocare. În plus, există câteva cerințe critice pe care un sistem de asistență a șoferului trebuie să le îndeplinească pentru a putea fi considerat o posibilă soluție de implementat la bordul unui vehicul: costul sistemului să fie suficient de scăzut pentru a permite încorporarea lui în orice vehicol produs în serie, sistemul să fie suficient de rapid încât să detecteze și să recunoască obstacolele în timp real; de asemenea, el trebuie să fie eficient (să detecteze toate obstacolele și foarte puține alarme false) și robust (să fie capabil să facă față la diferite condiții de mediu dificile).

Pentru sistemul propus s-au căutat senzori care să poată furniza suficientă informație pentru a detecta obstacolele (chiar și cele ocluzate) în orice situație de iluminare sau vreme, pentru a le recunoaște și a le identifica poziția în scenă. În domeniul vehiculelor inteligente nu există un astfel de senzor care să poată rezolva singur toate aceste cerințe. Există însă sisteme care folosesc unul sau mai mulți senzori diferiți pentru a realiza detecția, recunoașterea sau urmărirea obstacolelor, sau chiar combinații ale acestor funcții. După compararea avantajelor și a dezavantajelor tehnologiilor pasivă și activă, s-au ales senzorii cei mai potriviți pentru sistemul de detecție și recunoaștere a obstacolelor propus. Datorită posibilităților de interferențe care pot apărea între diferiți senzori activi, interferențe ce pot fi critice pentru un număr ridicat de vehicule ce se mișcă simultan în același mediu, ne-am concentrat doar asupra folosirii senzorilor pasivi, neinvazivi, adică a camerelor video. Astfel, sistemul propus folosește camere cu funcționare în spectrul vizibil și infraroșu, alese relativ complementare, deoarece sistemul trebuie să poată funcționa bine în diferite condiții dificile, precum iluminare slabă sau vreme proastă (cum ar fi întuneric, ceață sau ploaie).

Sistemele monomodale sunt adaptate la o singură modalitate, fie vizibilă fie infraroșie și chiar dacă ele furnizează rate de recunoaștere bune pe setul de test, aceste rezultate pot fi îmbunătățite și

mai mult prin procesarea combinată a informației vizibile și infraroșii, ceea ce înseamnă în cadrul unui sistem bimodal. Sistemele bimodale pot lua diferite forme în funcție de nivelul la care este combinată sau fuzionată informația. Astfel, propunem trei sisteme de fuziune diferite: la nivelul caracteristicilor, la nivelul kernelelor SVM, sau chiar la un nivel mai ridicat, la nivelul scorurilor de potrivire furnizate de SVM. Fiecare dintre aceste sisteme îmbunătățește performanțele comparativ cu sistemele monomodale. Pentru a asigura adaptarea sistemului la condițiile de mediu, în schemele de fuziune kernelele, scorurile de potrivire și caracteristicile au fost ponderate (cu un coeficient de ponderare al sensorului) în concordanță cu importanța relativă a senzorilor de modalitate. Aceasta permite obținerea unor performanțe mai ridicate la clasificare. În cadrul fuziunii scorurilor de potrivire de asemenea există posibilitatea de a realiza în mod dinamic adaptarea coeficientului de ponderare la context.

Pentru a reprezenta informația despre obstacolele ce trebuie recunoscute de sistemul de recunoaștere al obstacolelor, s-au preferat câteva tipuri de caracteristici pentru codarea informației existentă în imaginile obstacolelor. Aceste caracteristici sunt obținute în modulul de extragere al caracteristicilor și ele sunt: caracteristici wavelet, caracteristici statistice, coeficienții unor transformate și altele. În general, modulul de extragere al caracteristicilor este urmat de un modul de selecție a acestora, în care este estimată importanța lor și doar acele caracteristici care sunt cele mai relevante vor fi alese ulterior pentru a reprezenta informația. Diferite metode de selecție a caracteristicilor sunt testate și comparate pentru a evalua relevanța fiecărei caracteristici (și a fiecărei familii de caracteristici) raportat la obiectivul nostru de clasificare a obstacolelor. Relevanța fiecărui vector de caracteristici construit pe baza acestor metode de selecție a fost evaluată prima dată pe baza unui clasificator KNN (k Nearest Neighbours, în română - cei mai apropiați k vecini) (cu numărul de vecini  $k = 1$ ), datorită simplității acestuia la utilizare: el nu necesită un proces de optimizare a parametrilor așa cum necesită clasificatorul SVM.

Pentru a crește acuratețea clasificării, dar și pentru a obține un clasificator puternic, mai parametrizabil pentru schemele de fuziune propuse, clasificatorul KNN a fost ulterior (după ce a fost aleasă metoda cea mai bună de selecție a caracteristicilor pe setul de antrenare și cele mai relevante caracteristici au fost selectate) înlocuit cu un clasificator SVM. Deoarece nu se cunoaște dinainte ce combinație de hiper-parametrii ai SVM este cea mai potrivită pentru o anumită problemă de clasificare, a fost nevoie de o operație de căutare a modelului, realizată printr-o tehnică de validare încrucișată prin 10 directoare, care să furnizeze kernelul optimizat pentru SVM, kernel folosit ulterior pentru fiecare schemă de fuziune și pentru fiecare vector de caracteristici considerat.

În final, s-au testat schemele de extragere a caracteristicilor, de selecție a acestora și de fuziune propuse pentru o problemă cu 4 clase, adică s-a realizat discriminarea între vehicule, pietoni, cicliști și obstacole din fond. Rezultatele au demonstrat că toate sistemele bimodale vizibil-infraroșu sunt mai bune decât cele monomodale. Fuziunea este eficientă și robustă deoarece permite îmbunătățirea ratelor de recunoaștere. În plus, schema de selecție a caracteristicilor furnizează un vector cuprinzând doar cele mai relevante caracteristici pentru procesul de clasificare. Această reducere a dimensiunii vectorului de caracteristici pe lângă faptul că produce rate de acuratețe mai ridicate, permite și reducerea timpului de calcul care este crucial în acest tip de aplicații.

**Cuvinte cheie:** Fuziune, Cameră cu infraroșu, Extragere de caracteristici, Selecție de caracteristici, Mașina cu suport vector, Nucleu, Scoruri, Optimizare de hiper-parametrii, Căutare de model, Validare încrucișată cu 10 directoare.

---

## THESIS STRUCTURE

The research presented in this dissertation advances the theory, the design and the implementation of the proposed Obstacle Recognition component in the frame of an entire Obstacle Detection and Recognition (ODR) system. The proposed recognition component is designed for improving the road security by discriminating between different types of obstacles from the road and it is based on the fused information provided by visible spectrum and infrared spectrum cameras. The present work contains 169 bibliographical references and it is structured in five chapters as follows.

The **first chapter** is intended to give a motivation for why the ODR task is an important area to be investigated, and how the work done in this thesis can contribute to an ODR system. It also introduces the basic information necessary to understand the main characteristics and problems of the ODR task. The fundamental requirements for developing an affordable-price, real-time, efficient and robust system to be deployed on board of the vehicle are presented, followed by the specific characteristics of the ODR systems from the intelligent vehicle domain. Finally, our proposed solution and how we intend to solve all the specified requirements for the ODR system are introduced.

**Chapter 2** is dedicated to the systems (and the sensors they employed) from the intelligent vehicle field which addressed a problem similar to our, therefore it is a state of the art. Different types of sensors are investigated and their advantages and drawbacks are presented in the frame of most cited systems developed in the intelligent vehicle domain. The main types of sensors were examined, but we concentrated especially on the information each type of sensor could provide. Some sensors may have many advantages, but also some strong limitations, which make them to be not-so-properly for the implementation of an ODR system. The chapter is mainly focused on comparing advantages and disadvantages between the passive and active technologies and choosing the best solution for developing an ODR system. Considering the high price and the interference problems, we chose not to employ any active technology for the proposed system. In this chapter we motivated our choice to use only cameras, so passive sensors operating in a non-invasive way and which in addition are also cheaper than their counterparts, the active sensors.

In the next three chapters, our proposed system is presented.

In **Chapter 3** the baseline Obstacle Recognition component is presented, in the frame of an entire ODR system. The problems addressed here are intended to make a detailed presentation of the functioning mode and of the components that form this base system. The Obstacle Recognition component is more emphasized, and the following are also presented: the image database on which the proposed schemes have been experimented, the measures by which the performances of these schemes have been evaluated, but also how the feature vector that will characterize each instance within the system was composed. Basic notions about the classifier used in the frame of the developed fusion schemes, which is a SVM, are also presented. The individual or monomodal visible and infrared systems are also illustrated, together with a first set of experiments realized with these simple systems.

**Chapter 4** is structured in two main parts, the first one is presenting a motivation for why the step of features selection is needed and the main possibilities to accomplish this task are given. Different features selection algorithms are presented, tested and evaluated in order to compute the most pertinent feature vector to encode the information from the image database. Our method to perform the features selection is described and the last part is presenting the experiments we realised in order to perform the selection of features by the mentioned methods. Possible improvements are studied and implemented in order to choose the best feature vector to encode the information provided by the visible and infrared cameras. Once obtained, this feature vector could improve the accuracy of the system, but also it could decrease the processing time needed for the system in the Obstacle Recognition stage.



In **Chapter 5**, three different fusion schemes are presented and evaluated having the main purpose the improvement of the recognition accuracy, but also the possibility to adapt the system to different context situations. Fusion is performed at different levels, low or high (by combining features, respective matching scores), but also at an intermediate level: fusion at the kernel level, which is the solution we propose for our final system. In this last chapter the monomodal systems are also brought in discussion, but the main processing is done with bimodal systems, thus combining both visible and infrared information. They use the bimodal information at different stages, depending on the applied fusion scheme. A comparative study of individual visual and infrared obstacle recognizers versus fusion-based systems is performed and the obtained results are presented and discussed.

In the last chapter, we draw the main conclusion about the proposed fusion schemes and several potential improvements of our work are given.