On the track-to-track association problem in road environments

Nikos Floudas, Panagiotis Lytrivis, Aris Polychronopoulos, Angelos Amditis
Institute of Communications and Computer Systems (ICCS), Athens
Iroon Polytechniou St. 9, 15773 Athens, Greece,
{nikos.floudas,panagiotis.lytrivis,arisp,a.amditis}@iccs.gr

Abstract – Multi-sensor systems in automotive safety applications and sensor data fusion have become very popular in recent years. Sensors on board cars and active safety applications are increasing in number and the need to define a common method for object extraction and serving these applications has been recognized. Authors propose a high level fusion approach suitable for automotive sensor networks with complementary or/and redundant field of views. The advantage of this approach is that it ensures system modularity and allows benchmarking, as it does not permit feedbacks and loops inside the processing. In this paper this track level data fusion approach is introduced with the main focus to be on the data association of tracks coming from the on board sensors with distributed processing. The core of the proposed approach is the formulation of the data association problem in presence of multipoint objects and then the solution for (a) multidimensional assignment and (b) all around vehicle object maintenance. The motivation of this work is the research work that is carried out in the project PReVENT/ProFusion2 where the proposed algorithm is being tested in two experimental vehicles.

Keywords: data association, heterogeneous sources, track level fusion, automotive safety applications

1 Introduction

The motivation of this work is the research in the EU funded IP PReVENT [12] ProFusion2 subproject concerning the development of data fusion algorithms for object refinement. This includes the issue of managing information, from multiple sensors or sensor systems, in a common platform for advanced vehicle applications extracting high level information for typical objects of road environments. In these systems sensor fusion has been recognized as beneficial for their operation and functional extension.

There seems to be a wide recognition in intelligent vehicles engineering that both vision sensors and range sensors like radars and lasers are absolutely useful and essential. The only obvious solution towards simultaneously exploiting this multi-source and heterogeneous amount of information is data fusion. At the same time as more and more applications (i.e. forward collision warning, lane change aid, etc) are integrated in the same vehicle, more dedicated to applications sensors are deployed as well. This is one other reason why data fusion has to be applied; in order to handle this redundant and complementary information in a global manner, ensuring the extension of single sensor efficiency and making system design economic as a set of applications would share the same sensors [1].

However, the unanimous acceptance of sensor data fusion (SDF) has generated in turn a series of new arguments for which SDF architecture is the most appropriate for the problems in automotive area. The main architectures are the High Level Fusion (HLF) and the Low Level Fusion (LLF) approach. SDF in this work refers to object recognition only, with the HLF to imply a level of distributed data processing coming from the on board sensors with distributed processing. The core of the proposed approach is the formulation of the data association problem in presence of multipoint objects and then the solution for (a) multidimensional assignment and (b) all around vehicle object maintenance. The motivation of this work is the research work that is carried out in the project PReVENT/ProFusion2 where the proposed algorithm is being tested in two experimental vehicles.

Keywords: data association, heterogeneous sources, track level fusion, automotive safety applications

1 Introduction

The motivation of this work is the research in the EU funded IP PReVENT [12] ProFusion2 subproject concerning the development of data fusion algorithms for object refinement. This includes the issue of managing information, from multiple sensors or sensor systems, in a common platform for advanced vehicle applications extracting high level information for typical objects of road environments. In these systems sensor fusion has been recognized as beneficial for their operation and functional extension.

There seems to be a wide recognition in intelligent vehicles engineering that both vision sensors and range sensors like radars and lasers are absolutely useful and essential. The only obvious solution towards simultaneously exploiting this multi-source and heterogeneous amount of information is data fusion. At the same time as more and more applications (i.e. forward collision warning, lane change aid, etc) are integrated in the same vehicle, more dedicated to applications sensors are deployed as well. This is one other reason why data fusion has to be applied; in order to handle this redundant and complementary information in a global manner, ensuring the extension of single sensor efficiency and making system design economic as a set of applications would share the same sensors [1].

However, the unanimous acceptance of sensor data fusion (SDF) has generated in turn a series of new arguments for which SDF architecture is the most appropriate for the problems in automotive area. The main architectures are the High Level Fusion (HLF) and the Low Level Fusion (LLF) approach. SDF in this work refers to object recognition only, with the HLF to imply a level of distributed data processing in sensor level, while in LLF all the sensor data are processed centrally. Both approaches demonstrate a set of known well advantages and disadvantages (e.g. [2]), but the heterogeneous information of automotive systems changes significantly the situation. Authors are in favor of HLF as this approach offers modularity and allows benchmarking and substitution of sensorial systems as it does not permit feedbacks and loops inside the processing. Consequently, HLF requires less computational load and communication resources compared to LLF. Yet, in HLF the issue that some times arises is that the implied modeling of sensor level information is vague and even contradictory with the assumptions in the fusion processor [3], and certainly that the sensor level errors are depended with each other and should be taken into consideration. However, these limitations could be overcome, taking advantage at the same time of the many benefits of this scheme.

In this paper a specific generic approach of HLF architecture is proposed that takes the information provided by sensorial systems and provides as output the global fused objects describing the vehicle environment. The architectural level where this operation takes place is the intermediate level between the measurements and the applications, mentioned often as the perception layer. This approach is referred in the next as: Track Level Fusion (TLF).

Related work on multi sensor data fusion for preventive safety has been carried out in a series of research activities such as the ARCOS project [4] with a forward collision mitigation combining stereo vision and laser scanner, the EUCLIDE research project [5] which is
a forward collision warning and vision enhancement application using a far infrared camera and a millimeter wave radar sensor. The same sensor combination was used in the PAROTO project [6], while in the CARSENSE [7] project information from a radar, video sensors and a laser was fused. These first fusion systems in automotive safety were using a limited number of sensors and were focused in one particular application (e.g. forward collision warning). However, in IP PReVENT more sophisticated fusion systems are being tested with a group of integrated applications, characteristic example is the LATERAL SAFE subproject with eleven sensors observing holistically the rear and side areas of the vehicle [8].

In ProFusion2 innovative research on SDF takes place and prominent car manufacturers offer useful test cases where several independent sets of sensorial systems are used. This approach is being tested in two specific prototype vehicles: (a) a Volvo Technology Corporation sensor equipped truck with main test case the multidimensional assignment, (b) a Centro Richerche FIAT passenger car with all around sensor coverage. More details on these test vehicles are given in the next section.

The structure of the paper is the following: Section 2 presents the two test vehicles the objectives and requirements of the fusion approach and their innovative aspects. Section 3 describes the fusion algorithm proposed from an architectural point of view and Section 4 gives a brief overview of its design. Some first results of the development of these methods concerning sensor level processing and data association are present in Section 5. Finally Section 6 closes the paper.

2 Objectives and test cases

In this section the main objectives and requirements of the fusion approach described in this work are given, together with the description of the sensor systems of the two test cases of the algorithm. The section closes with the innovative aspects of this fusion approach.

2.1 Requirements

To begin with the main objective of sensor data fusion in automotive technology is to serve the active safety applications by providing a consistent and trustworthy list of detected objects in the vehicle’s environment. The information should include position and velocity estimates, motion dynamics information, ID based history of the objects and if possible information on objects dimensions and some kind of classification index (e.g. passenger car, truck, motorcycle etc). Certainly, this additional information is heavily depended on the type and the quality of the available sensors. It is always true that each specific application requires a specific type of sensor and dedicated signal processing, but in ProFusion2 it is the first time an effort is made to tune a global algorithm proper for handling several sensors input and supplying different kinds of applications.

Coming to the specific track level fusion architecture it should be pointed out that reliable and consistent signal processing/tracking in the single sensor level is required, as far as estimation, maintenance of tracks and object extraction is concerned. This is the focus of dedicated processing modules applied in sensor level data with two sub cases that can be identified: (a) if the sensor is delivering raw data then a complete tracking algorithm will be applied, (b) if a level of processing is performed in sensor level then most possibly only a clustering/grouping of sensor output for object extraction is usually performed; but in either case this depends on the type of sensor and the sensor specific output adjustment needed.

As far as the track to track fusion is concerned it should be adaptable to sensor number, type and topologies and it should eliminate the transfer of sensor processing errors to the fusion level. Moreover, the overall algorithm is expected to run within a “sensible” time duration and computational load.

2.2 Test vehicles

As already mentioned the designed and developed fusion algorithm will be integrated and tested in two test vehicles.

The Volvo Technology Corporation sensor equipped truck is the main test case for the multidimensional assignment that comprises sensors with redundant field of view looking forward: a long-range radar (LRR), a laser scanner, two short-range radar (SRR) sensors, and a lane camera system. The FOVs of these sensors together with their exact location and coverage are shown in Fig. 1. The main functionality tested in this vehicle is the collision mitigation by braking and the system aims to combine the information from different sources about the same object to better derive its properties and object classification.

![Fig. 1: Sensor coverage in the test truck](image)

The other test vehicle is a sensor equipped passenger car by Centro Richerche FIAT that hosts a set of applications, namely collision mitigation, pre-crash detection, lane change aid, lateral collision warning and lateral and rear monitoring. All these applications require an all round sensor coverage and data fusion of a great number of on board sensors: six side looking short range radars, one rear looking long range radar, one forward looking long range radar, two forward looking short range radars, one...
forward looking camera and two rear looking cameras on the two mirrors. Sensor coverage is plotted in Fig. 2.

Fig. 2: Sensor coverage in the test car

2.3 Innovation

Within ProFusion2 the issue of novelty of SDF in automotive safety systems is highly stressed. In this paragraph a short overview of the innovative aspects of our approach is given. These comprise four main issues: (a) management of multi-source heterogeneous information (many and of different type sensors) in a moving platform with a generic global algorithm applicable for various sensor topologies and selections, (b) solution of new data association problems for single and multipoint objects together with multidimensional assignment (more than two sensors), (c) the all-around platform vehicle tracking and maintenance of other moving vehicles and (d) the development of a sensor tracking algorithm that deals with various topologies (e.g. rear or side looking sensors) and various types of sensors.

3 Fusion System Architecture

As mentioned in the introduction authors support the HLF approaches for managing the multi-source and significantly heterogeneous information of multi-sensor equipped vehicles. Apart from the known issues of common sensor modeling, depended estimation errors in the sensors level, HLF concerns also the topics of filtering and state estimation, having to deal with the complicated issues of data association that arise in such systems. These issues are discussed in sections 4 and 5 in more detail, with a brief overview of their architectures to be given in this section.

The main parts of the TLF algorithm as illustrated in Fig. 3 are: the time and space alignment of track arrays, the division of fusion sub problems according to the area covered by each sensor or sensor system, the track to track association procedure that is solved with 2D and S-D (with $S \geq 3$) assignment, the fusion object update from the pairs or S-ples of tracks and the object management that is the final step before the objects pass to the output. As it is mentioned input track arrays could be the output of a dedicated tracking or other processing module.

Fig. 3: Track-Level Fusion architecture

The core of Track Level Fusion (TLF) is the track to track association algorithm. This plays a key role in the performance of TLF ensuring the continuity and maintenance of objects all around sensor covered area and the solution of multi-source objects assignment. Important part of TLF approach is the Sensor Tracking (ST) of single sensor measurements. The output of this procedure is the high level track information input to the TLF system illustrated in Fig. 3. ST architecture is shown in Fig. 4, the main characteristic of the algorithm that is developed is the capability with simple functionalities to deal with different sensors’ input (type, measurement, orientation) providing the best possible solution. ST algorithms applicable for different types of radars and laser scanner and vision systems are being developed.

Fig. 4: Sensor level tracking architecture

4 Track Level Fusion Approach

As written in Section 3 the key components of the TLF approach are the ST algorithm for the set of the available sensorial systems and data association module that should deal with all around object maintenance and multidimensional assignment.

4.1 Mathematic formulation

Skipping the well known equations of state estimation, we present here the formulations of the multidimensional data association problem. Suppose we have $N$ data sources applicable for association with $M^p$ observed values from each source $p$ with $p=1,2,\ldots,N$. Next the following quantity is defined $z_{i,j}^{p}$, which corresponds to the hypotheses of association formulations, where
observations $i_1, i_2, ..., i_N$ come from the same target-source. For instance $z_{322}$ refers to the fact that observation 3 of source 1, observation 2 of source 2 and observation 2 of source 3 come from the same target. If any index is equal to zero means that this source gives no detection. Thus the dual variable $z_{i_1i_2...i_N}$ for the association of a hypothesis is defined as:

$$z_{i_1i_2...i_N} = 1, \text{ track hypothesis is correct}$$

$$z_{i_1i_2...i_N} = 0, \text{ track hypothesis is false}$$

In a similar manner the cost of formation of associations $c_{i_1i_2...i_N}$ is defined. The prediction that the observation of source $p$ is a false alarm has the following cost $c_{0...i_p...0} = 0$. Taking all these definitions into consideration the problem of generation of associations using data from $N$ sources is transformed into the subsequent optimization problem:

$$\max \sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} c_{i_1i_2} z_{i_1i_2} \quad (1)$$

given that:

$$\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} z_{i_1i_2} = 1, \quad i_1 = 1,2,...,M_1 \quad (2.1)$$

$$\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} z_{i_1i_2} = 1, \quad i_2 = 1,2,...,M_2 \quad (2.2)$$

$$\vdots$$

$$\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} z_{i_1i_2} = 1, \quad i_N = 1,2,...,M_N \quad (2.N)$$

All the above constraints could be expressed as:

$$\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} \sum_{i_3=0}^{M_3} \sum_{i_N=0}^{M_N} z_{i_1i_2...i_N} = 1, \quad \forall i_p = 1,2,...,M_p, \quad \forall p = 1,2,...,N \quad (3)$$

Equation (3) shows that all observations of source $p$ should be taken into account only once in order to produce all possible combinations of the rest of the sources. The case that the associated sources are two is solved by known constrained optimization problems solution algorithms (e.g. auction algorithm) for 1-to-1 assignment, and with possible extension to probabilistic solutions. The not optimum but efficient solution to the multidimensional assignment problem is feasible with the method of Lagrange multipliers relaxation.

**Track Fusion** of data of distributed sensors in HLF architectures is the next step in TLF approach. In TLF a set of $N$ different sensors with $M_n$ tracks from each one, after the generation of the associated groupings, let them be $G$ of $M_G$ tracks of each, finally gives the $G$ ultimate fused objects together with the objects observed by solely one sensor. A track consists, apart from the estimated state vector, also of a quality measure, which usually is the covariance matrix of the estimation error. The track array coming from sensor $n$ ($n$ from 1 to $N$) is: $\{x_{1n}, P_{1n}\}, \{x_{2n}, P_{2n}\}, ..., \{x_{Mn}, P_{Mn}\}$

What is sought for each group of associated sensor level tracks is a fused track (object) that is a best estimation compared to each of the single sensors outputs individually. The estimation of the state and covariance of the fused object $i$ (with $i$ between 1 and $G$) that arrives from estimations of $M_G$ sensors, specified as $\lambda_1, \lambda_2, ..., \lambda_{MG}$ ($M_G$ from 1 to $N$) are functions of these parameters:

$$x_{\text{fused}} = f_1(x_{A1}, P_{A1}, x_{A2}, P_{A2}, ..., x_{A_{MG}}, P_{A_{MG}}) \quad (4)$$

$$P_{\text{fused}} = f_2(P_{A1}, P_{A2}, ..., P_{A_{MG}}) \quad (5)$$

There are several approaches to find the most appropriate method to identify the best functions $f_1$ and $f_2$ in order for the fusion requirements to be met.

### 4.2 System design

The several steps of design process of the TLF approach require firstly a robust ST algorithm for the single sensor data. Then the main fusion algorithm comes which is heavily based in the DA performance.

Therefore, one ST applicable to the different available sensors long range radar and short range radars, laser scanner and vision systems is developed. Moreover different topologies for each of them are applied. Processing of vision systems (e.g. image processing) does not take place and data are taken from built-in systems. Sensor specific ST algorithms that take account of different object measurements, measurement models, object occlusions with small modifications are under development. It should be pointed out that sensor specific object extraction is required in order to formulate the data association problem.
Two general categories (Fig. 5) of assignment problems are identified: the classical 2D assignment problem and the S-D (with \( S \geq 3 \)). The first is most common in the problems in the typical sensor topologies in automotive area but the second can be also observed in the cases of more than two sensors observing a common area, as happens in the Volvo test vehicle in ProFusion2.

Auction and JVC algorithms and their switching selection are adequate to solve the 1 to 1 assignment problem in 2D data association. 2D data association is completed with the integration of the case of 1 to N assignment [9] to the overall algorithm. The case of three or more sensors observing a common area was also investigated. The typical Lagrangian relaxation method is used to solve this multidimensional data association case. The process of sequential relaxation of constraints and reduction in subproblems of lower dimension and then the Lagrangian multipliers update phase until an assignment solution will be found, is illustrated in Fig. 6.

5 First results

In this section some first preliminary results towards the system development are presented. These include results of single sensor processing and first simulation and development of the TLF system.

5.1 Results on object extraction

Here the first outcome of object extraction from different types and orientations of customary sensors in automotive technology is briefly discussed. As far as the LRR is concerned it should be mentioned that one or two maximum detections per objects is the normal case (Fig. 7) while the SRR in general gives more detections per actual object (Fig. 8), especially if the object is detected from the side area. The laser scanner output comes from a dedicated module that gives already a set of polygons for the objects detected (example in Fig. 9). Examples of these types of detections with cumulative plots of data are given in the following figures. A general idea of typical returns of these sensors can be drawn from the following three figures; sensor specific object extraction is being investigated. Those figures contain a cumulative plot of the sensors measurements for the same number of successive scans from each sensor individually and for the same road traffic scenario.

![Fig. 7: Cumulative plot of LRR data (one object)](image)

![Fig. 8: Cumulative plot of SRR data (two objects)](image)

![Fig. 9: Cumulative plot of laser scanner data (objects and road borders)](image)
in finding the solution to the 2D assignment problem for different sizes of association matrices and a simulation study of a three sensor data fusion (TLF) system concerning the performance of data association (object recognition, number of missing objects and false alarms).

5.2.1 Time delay of 2D data association

2D assignment is solved by typical constrained optimisation solution algorithms like auction algorithm [10] and JVC algorithm [11]. Both these two solutions have been implemented and tested and a comparison study was performed. The results comply with what is also known in literature (e.g. [11]) that auction algorithm is preferable (faster) in sparse matrices while JVC is faster in dense matrices. What can be drawn from this is that a scheme that firstly checks the sparsity of the association matrix and then selects between JVC and auction which algorithm should be followed.

Fig. 10: Comparison of execution time vs. 10x10 matrix sparsity

Fig. 11: Comparison of execution time vs. 40x40 matrix sparsity

An example of this testing for random 10x10 assignment matrices for these two algorithms is given in Figure 10. In x-axis the level of sparsity in a matrix is the percentage of non-zero elements in it. In y-axis there is the normalized mean execution time for each algorithm. The auction algorithm is faster than JVC only in very sparse matrices e.g. non zero elements less than 30%. This is more obvious when the size of matrices is bigger. In small matrices, like the 10x10 shown in Fig. 10 which are closer to the size of the matrices appearing in the problems we investigate, the time delay is not so obviously different as both algorithms seem to converge quickly in the solution. In the case of 40x40 matrices it is explicit that for non zero elements more than 40% in the matrix the JVC algorithm converges faster compared to auction, and also that the time delay of auction algorithm increases significantly with the increase of matrix density.

5.2.2 Performance of multidimensional data association – simulation

Simulations of various multidimensional tracking and fusion systems have been implemented. The Lagrangian relaxation method was applied in a tracking system of three sensors in a TLF architecture. Various scenarios of one or more moving targets were considered (first column of Table 1). At the same time each sensor tracked a random set of possible targets. For these movement scenarios the corresponding algorithms of the TLF were executed, and the correct operation of data association method was verified.

Table 1 summarises the TLF testing concerning the performance of 3D data association algorithms. The first column shortly describes the investigated scenario including the number of true targets and second column has the total number of scans of each scenario; these two columns show the total number of expected objects. The next three columns include the simulated measured objects by each of the three sensors (excessive amount of clutter was intentionally added to check the efficiency of data association). The following three columns include the results of sensor level tracking where the clutter is to a great extent reduced by the internal 2D assignment and
track management algorithms. The final column includes the results of the final fusion step where the implementation of the 3D Lagrangian relaxation assignment algorithm takes place. In general in most of the scenarios the fused objects are almost the same with the true apart from one scenario where some false objects remain. This testing demonstrates in general the satisfactory performance of the data association algorithm; however it should be pointed out that more testing with actual data recordings is necessary.

5.3 Results on fusion

The first input data from system integration are already available and preliminary testing is being performed. One screen shot of the offline evaluation tool is given in Fig. 12. In Fig. 12 there is a synchronized video image of a road traffic scenario with the also synchronized data of the sensors to be plotted as well with different colors. An indication of the measurements velocity is also given. More analytic results of processing with real data would be available in the next months, concerning sensor level tracking and track level fusion.

6 Conclusion

In this paper a short overview of current research work of the authors in automotive safety area in SDF is given. The TLF generic fusion approach is proposed that its ambitious goal is the operation of reliable object extraction irrespective of sensor types and topologies and aiming at supporting simultaneously a set of various active safety applications. Some first results on modular level and specifically on the data association issue are also given.

Acknowledgements

The authors would like to thank the ProFusion2 consortium for the excellent cooperation and particularly Volvo Technology and CRF for the test vehicle system integration and the data recordings.

References


Fig. 12: Offline processing tool for evaluation of fusion algorithm

TS-LRR = 4016.13s
TS-SR1 = 4016.0957s
TS-SR4 = 4016.0957s
TS-LSC = 4006.1322s

Video

LRR sensor objects = 5
SRR-1 sensor objects = 2
SRR-4 sensor objects = 2
Laser Scanner objects = 11