Cloud Robotics Paradigm for Enhanced Navigation of Autonomous Vehicles in Real World Industrial Applications

Elena Cardarelli, Lorenzo Sabattini, Cristian Secchi and Cesare Fantuzzi

Abstract—Autonomous vehicles require advances sensing technologies, in order to be able to safely share the environment with human operators. Those sensing technologies are in fact necessary for identifying the presence of unforeseen objects, and measuring their position and velocity. Furthermore, classification is necessary for effectively predicting their behavior. In this paper we consider the presence of sensing systems both on-board each vehicle, and installed on infrastructural elements. While the simultaneous presence of multiple sources of information heavily improves the amount (and quality) of available data, it generates the need for effective data fusion and storage systems. Hence, we introduce a centralized cloud service, that is in charge of receiving and merging data acquired by different sensing systems. Those data are then distributed to the autonomous vehicles, that exploit them for implementing advanced navigation strategies. The proposed methodology is validated in a real industrial environment to safely perform obstacle avoidance with an autonomously driven forklift.

I. INTRODUCTION

This paper deals with advanced sensing systems and data fusion techniques for Automated Guided Vehicles (AGVs) used for automated factory logistics, that is for the transportation of raw materials or final products in industrial environments.

Regardless their specific application, automated vehicles require appropriate exteroceptive sensing systems, that make them able to provide reliable obstacle detection, gathering information about the environment. Several sensing technologies have been employed in the obstacle detection field [1], such as computer vision [2], [3], radar [4]–[6] and lidar [7]–[9].

Different kinds of sensors provide complementary information regarding the entities that are in the neighborhood of the vehicle. In order to obtain a consistent, robust and accurate representation of the environment, it is necessary to combine the information coming from multiple sources. This process is generally referred to as data fusion: review and discussion of several data fusion definitions is presented in [10].

In particular, significant attention has been dedicated to the data fusion in recent years: a review of contemporary data fusion methodologies, as well as the most recent developments and emerging trends in the research field is presented in [11]. As discussed in [12], data fusion techniques can be classified based on the abstraction level under consideration. In particular, we can distinguish the following classes:

- **Low level** fusion: this level deals directly with raw data to improve the accuracy of the individual sources.
- **Medium level** fusion: it is based on the processing of features or characteristics (i.e. dimension, shape, position). This level is also referred to as the feature or characteristic level.
- **High level** fusion: also known as decision fusion, this level addresses symbolic representation, such as object classes.
- **Multiple level** fusion: based on simultaneous processing of data provided at different levels of abstraction.

According to point where data fusion is performed it is possible to distinguish three main types of classification architecture: distributed, centralized and hierarchical. In a distributed architecture each source node processes raw data independently, thus computing an estimation of the properties of the acquired objects based only on local views, which does not guarantee a complete knowledge of the environment. Conversely, in a centralized architecture a single computation unit gathers data from all the input sources, and performs fusion of the received data. While this architecture provides a complete knowledge of the environment, due to multiple points of view, the main drawbacks of this solution are related to the large amount of bandwidth requested to transmit raw data over the network. A tradeoff between these methods is represented by hierarchical architectures, where fusion process is performed at different levels in the hierarchy. Namely, local nodes process raw data locally, and then provide the central computation unit with the results of the local processing. This leads to avoiding the necessity of transmitting large amounts of raw data, while still guaranteeing a global view of the environment.

Multi-sensor data fusion is therefore a multidisciplinary technology that involved several application domains, such as robotics [13], [14], military application [15], biomedicine [16], [17], wireless sensor network (WSN) domain [18], video and image processing [19]. In industrial environments, such as in highly automated logistics systems [20]–[23], the employing of multi-sensor data fusion is not a widespread solution, despite the presence of autonomously driven forklifts equipped with several sensor technologies.

The contribution of this paper is in the definition of a multi-sensor and data fusion system for enhanced navigation of autonomous vehicles in industrial environment, shared with human operators. In particular, we introduce a multiple...
level data fusion techniques, implemented on a centralized cloud system, that is in charge of gathering information from the sensors, and sharing the output of the data fusion process with the autonomous vehicles.

The paper is organized as follows. Section II describes the scenario under consideration. Subsequently, in Section III we introduce the data fusion cloud system, and the multiple level data fusion technique, whose experimental validation is described in Section IV. A case study is analyzed in Section V, where the proposed methodology is experimentally validated for enhanced navigation of an AGV in a real world industrial environment. Finally, Section VI contains some concluding remarks.

II. SCENARIO

AGVs are nowadays gaining popularity for logistics applications. The most well known example is represented by Kiva Systems [20]–[22], whose solution provides a completely automated AGV system for goods delivery in a distribution center. Complete automation is however not achievable when dealing with logistics facilities within production plants, where heavy parts need to be handled [23], [24]. In those cases, it is mandatory for the AGVs to share the working environment with human operators and manually driven forklifts. Therefore, safety is the a major issue to be taken into account, and is generally managed by means of certified laser scanners. While those sensors are very reliable in identifying the presence of obstacles in the neighborhood of the AGVs, they are unable to classify the detected objects.

Hence, severe constraints need to be applied on the motion of the AGVs, in order to avoid safety issues related to unpredictable events. As illustrated in [25], in the vast majority of modern automatic warehouses, AGVs are constrained to move along a set of (virtual) paths, that is usually referred to as roadmap. Constraining the AGVs to move along a finite set of roads makes their behavior absolutely predictable, which facilitates the adoption of safety measurements. At the same time, this constraint severely reduces the flexibility of the system. In particular, this reduced flexibility clearly affects the performance of the system in the presence of unforeseen obstacles. In fact, if an obstacle suddenly appears in front of an AGV, it is necessary to re-plan the AGV’s path, in order to avoid collisions with the obstacle. If AGVs are constrained in the roadmap, re-planning means finding an alternative path on the roadmap, which is not always feasible: consider, for instance, the frequent case of mono-directional roads. In this case, if an alternative path cannot be found on the roadmap, the AGV gets stuck, until the obstacle has been removed.

Deviations from the roadmap are generally not allowed, for different reasons. The first one is the aforementioned lack of classification capabilities in commonly adopted safety laser scanners. The second reason is related to the fact that sensor systems installed on board each AGV are not able to acquire global information about the surrounding environment. Roughly speaking, they can not look around corners, or look behind obstacles. This issue is generally solved, for pedestrians and manually guided vehicles, exploiting systems installed at critical points in the infrastructures, such as hemispherical mirrors. This concept can be implemented for automated systems as well, installing perception systems (such as cameras or laser scanners) in critical points in the infrastructure, such as in the intersections among corridors [26]. Sensors on the infrastructure provide a complementary source of information, that enhances the onboard perception capabilities.

III. CLOUD SYSTEM FOR DATA FUSION AND STORAGE

In this section we introduce a hierarchical data fusion technique, implemented as a cloud system in the industrial environment.

Data fusion is necessary because of the presence of different sensing systems that simultaneously acquire data, which need to be made available to the AGV control system, for inclusion into the planning and control strategy. Therefore, we introduce a centralized system, that is in charge of receiving data from different sources, opportunistically merging them, and making them available for the AGV control system. This centralized system defines a Global Live View of the environment, that contains constantly updated information regarding all the entities that populate the industrial environment [27]. As a motivating example, consider the scenario depicted in Fig. 1. In this example, an AGV is in the presence of multiple objects, as well as pedestrians. Based only on local sensing (Fig. 1(b)), the AGV is able of identifying only a limited portion of the objects in its neighborhood. Conversely, if exploiting the Global Live View cloud service (Fig. 1(c)), the AGV is provided with global information, that integrates data acquired by different perception systems.

More specifically, in the proposed architecture the principal perception systems consist of infrastructure and onboard sensors with object detection, tracking and classification capabilities. Moreover, a static three-dimensional map of the environment is available in order to describe all the static infrastructural elements (e.g. rack, walls, doors, etc.) [28].

Thus, in the proposed architecture, the information about the obstacles in the scene may be provided by several sources, involving the possibility of data redundancy, inconsistency, ambiguity, noise, and incompleteness. To overcome this problem, the Global Live View collects all data acquired by the sensors and combines them in a unique and complete representation of the overall system, including the static and dynamic entities that act inside it. In particular, the Global Live View allows to achieving higher quality information, providing a global updated map representing the static entities (the 3D map of the plant, the roadmap), the dynamic entities (the current position and velocity of the AGVs, the position and velocity of currently identified objects), the congestion zones and the status of the monitored intersections.

Generally speaking, the information acquired by the infrastructure and on-board perception systems consists of tracked and classified objects, identified with a unique ID. In detail, data regarding each object are:
- Position, orientation, velocity, size.
- Class of the objects: human, manual forklift, AGV, other
dynamic object, static object.
- An assessment regarding the quality and reliability of
the classification.

The Global Live View is then updated with the information
acquired during the operation and a real-time global map is
generated. This output is shared with the AGV fleet in order
to improve their local on-board navigation capabilities and
support safe operation.

It is important to guarantee consistency with respect to
the real world: each virtual object represented in the map
must have a correspondence to a real object of the world.
Therefore, the Global Live View performs data fusion to
merge data acquired from the different sensors, reducing
information redundancy and verifying the presence of data
inconsistency and ambiguity.

Data fusion is a very well known problem, that has been
extensively studied in the literature. However, it is worth
noting that typical solutions consist in the fusion of low
level data (images, 3D point clouds, laser raw data). This is
however not practical for the application we are considering:
in fact, for each obstacle candidate, we assume to process
medium level features (ID, age, position, orientation, velocity
and size) and high level features (class and classification
quality) in order to optimize the data transmission time
and reduce the network overhead. Therefore, we propose a
two level methodology, that implements, separately, medium
level and decision level data fusion.

A. Medium level

In the described architecture, dealing data with fusion at
medium level means processing the object measurements
(ID, age, position, orientation, velocity and size) estimated
with uncertainty by the on-board and infrastructure systems,
as well as the elements inside the static map of the environ-
ment.

Therefore, from a medium level point of view we in-
troduce a heuristic, whose principal steps are described in
Algorithm 1. Since all positions, velocities and angles are
measured with respect to the local sensor coordinate system,
firstly, it is necessary to align the input lists of objects
coming from the sensors (i.e. obj_list_from_infrastructure and
obj_list_from_onboard), converting their data according to a
global reference system (see lines 3 and 5 in Algorithm 1).
Then, the algorithm calculates the occupational overlapping
among the different lists of objects detected by the infra-
structure and on-board perception systems, processing the
position and orientation of their bounding boxes (line 9).
Thus, a preliminary list of fused objects (line 10) is obtained
from the evaluation of the overlapping results, discriminating
among bounding boxes that represent different views of the
same object or separated elements. Through the integration of
the information about the velocities and directions estimated
for each fused obstacle (see procedure in line 11), it is then
possible to discriminate among static and dynamic obstacles.
Finally, the information representing the fused obstacles is
integrated in a grid map on which free space and unknown
regions are modeled, supporting the implementation of path
planning and navigation functions (details will be provided
in Section V).

B. High level

The high level data fusion for the implementation of the
Global Live View is tackled as a classifier combination
problem: the main steps of the algorithm designed to cope with this issue are defined in Algorithm 2.

In details, a preliminary classification result is determined for each object of the list obtained by the medium level data fusion (Algorithm 1), exploiting motion, directional and dimensional features (lines 3, 4, 5 and 6). These rough class values are then combined with the classification results provided by the sensors: the static 3D map of the environment, the on-board sensor systems and the infrastructure perception systems are considered as a set of classifiers that, given an input pattern, provide an output score that represents a confidence measure for each possible class of the system (human, manual forklift, AGV, other dynamic object, static object) to be the correct class for the input pattern.

Algorithm 2 High level data fusion

Require: obj_fused_list

1: procedure PRE-CLASSIFICATION
2:   for all fused_obj ∈ obj_fused_list do
3:     if fused_obj is static then
4:       Evaluate(size(fused_obj))
5:     else
6:       Evaluate(size(fused_obj), velocity(fused_obj))
7:   return pre_classified_list

8: procedure CLASSIFICATION
9:   for all pre_classified_obj ∈ pre_classified_list do
10:  Calculate Class Votes Average(pre_classified_obj)
11: return classified_list

Several methods can be found in the literature for solving the problem of classifier combination at measurement level (or type III [29]). Among these methods, we propose to exploit, in line 10, the average-rule aggregation scheme: namely, for each identified object, the average of the scores assigned to each class is computed, and the output is defined as the class with the highest average vote. Despite its simplicity, this elementary combination rule competes with the more sophisticated combination methods, as highlighted in [30], without requiring any training procedure. Moreover, this methodology is well suited for real time implementation, which is mandatory in this kind of application.

IV. DATA FUSION EXPERIMENTAL EVALUATION

The proposed data fusion strategy was evaluated in a real industrial environment, namely an automated warehouse where AGVs are exploited for logistics operations, whose map is represented in Fig. 2. In this experiment we utilized two sensing systems:

- an AGV is equipped with on-board exteroceptive sensors (namely laser scanners and cameras),
- a laser scanner mounted is on the infrastructure, in particular on a corner of the warehouse.

Both sensing systems were then exploited to acquire information on dynamic entities in the environment. The results of the experimental validation are shown in Figs. 3 and 4.

Data acquired from both sensing systems are locally pre-processed: the centralized data fusion cloud systems receives the position and the sizes of the bounding boxes of the acquired objects. In particular, Fig. 3(a) shows the bounding boxes of the objects acquired by the infrastructure sensing system, while Fig. 3(b) shows the bounding boxes of the objects acquired by the on-board sensing system. Data alignment is then performed, to harmonize the reference frames of the data acquired by the two sensing systems (Fig. 4(a)). The output of the data fusion algorithm is then represented in Fig. 4(b). As expected, data acquired by the two sensing systems are merged based on geometrical overlaps, whose analysis leads to reducing the outliers.

Further results are shown in the accompanying video.

V. CASE STUDY: ENHANCED AGV NAVIGATION

The proposed data fusion cloud service provides the AGV control system with constantly updated rich information regarding the entities in the environment. This information can then be exploited for implementing advanced control strategies for enhancing the efficiency of the system.

As a motivating example, we will hereafter consider the following case study: consider an AGV performing an assigned mission, thus traveling along a predefined path, that detects the presence of an obstacle that obstructs its motion.

It is worth noting that, considering only a local knowledge of the environment, based on safety laser scanners only, this situation is generally solved by halting the AGV, until the obstacle has been removed. In fact, in order to implement a local deviation from the roadmap in a safe manner, it is necessary for the AGV to:

- Have a reliable classification of the obstacle: in fact, while driving around a fixed object can be easily performed, the same cannot be done for overtaking a human operator, since human behavior cannot be predicted in a sufficiently reliable manner.
- Have a global knowledge of the surrounding environment: deviations from the roadmap should be avoided, if other dynamic entities (other AGVs, manually driven vehicles, pedestrians) are in the neighborhood of the current AGV’s position.
Local deviation from the roadmap can then be computed as detailed in [31].

To deal with the computation of a safe local deviation from the predefined path of the AGVs a specific module, called *Local Path Planning*, is introduced in the architecture, as represented in the UML sequence diagram in Fig. 5.

In particular, once the presence of an obstacle has been detected by the safety laser scanners on-board the AGV, a request for computation of local deviation is sent by the AGV to the Local Path Planning. This module contacts the Global Live View in order to receive information about the list of objects in the neighborhood of the AGV. These data are exploited by the Local Path Planning in order to assess the feasibility of a safe deviation maneuver, computing the curve for implementing the local deviation and transmitting it toward the AGV.

Since the Global Live View acts as a cloud system, enhancing information in the neighborhood of the AGV, the deviation maneuver is determined in a safe manner.

Experimental validation was carried out in an industrial environment, with an AGV controlled to overtake a fixed obstacle on the ground, based on information from the cloud data fusion system. The results of the experiments are shown in the accompanying video.

**VI. CONCLUSIONS**

In this paper we introduced a multi-sensor and data fusion system for industrial applications, where autonomous vehicles share the environment with human operators. In particular, we consider the presence of multiple sensing systems, installed both on-board each autonomous vehicle, and on specific infrastructural elements. While on-board sensing is necessary for ensuring safety, it provides only a local view of the neighborhood of the autonomous vehicle. Hence, additional sensing systems on the infrastructure complement this information, extending the field of view.
However, while the simultaneous presence of multiple sources of information heavily improves the amount (and quality) of available data, it generates the need for effective data fusion and storage systems. Thus, a centralized cloud system was introduced, that is in charge of gathering information from the different sensing systems, and implementing a multiple level data fusion strategy. The output of the data fusion process is then distributed to the autonomous vehicles, that can benefit from this enhanced sensing information for safely implementing advanced navigation strategies. The proposed methodology was validated in a real industrial environment, with an automated forklift driven to safely perform obstacle avoidance.

Current work aims at refining the medium level and high level data fusion strategies based on extensive evaluation on different real industrial scenarios, in the presence of multiple autonomous vehicles, and different dynamic entities. This will make it possible to utilize more advanced data fusion techniques based on training procedures.

References


