A Reinforcement Learning Modular Control Architecture for Fully Automated Vehicles

Jorge Villagrá, Vicente Milanés, Joshué Pérez, Jorge Godoy, Enrique Onieva, Javier Alonso, Carlos González, Teresa de Pedro, and Ricardo Garcia*

Autopia Program, Center for Automation and Robotics, UPM-CSIC
Ctra. Campo Real km. 0.200, 28500 Arganda del Rey (Madrid), Spain
{jorge.villagra,vicente.milanes,joshue.perez,jorge.godoy,
enrique.onieva,javier.alonso,carlos.gonzalez,teresa.pedro,
ricardo.garcia}@car.upm-csic.es

Abstract. This paper proposes a modular and generic architecture to deal with Global Chassis Control. Reinforcement learning is coupled with intelligent PID controllers and an optimal tire effort allocation algorithm to obtain a general, robust, adaptable, efficient and safe control architecture for any kind of automated wheeled vehicle.

1 Introduction

In the last years a huge effort has been made both in industrial and academic environments to develop advanced driving assistance systems (ADAS). As a result of this, several embedded control systems are already on board commercial vehicles: Adaptive Cruise Control, Lane Keeping or automatic Parking Assist systems. Most of these systems seek to solve a specific problem by considering a particular set of car actuators -active suspensions, brakes, traction torques and steering angles of each wheel. However, since the number of ADAS is significantly increasing, a global and modular control architecture should be considered to rationalize and, if necessary priorize, the use of all these actuators. Moreover, since energetic efficiency is nowadays one of the most challenging issues for any Intelligent Transportation System (ITS), this architecture should be oriented towards an optimal use of vehicle resources.

In the past, works like [1], [4] have proposed different variants of a Global Chassis Control (GCC). However, in all these cases the proposed architecture considered only security functions (ESP or ABS like systems). To the best of our knowledge, the only commercial system that introduces a general framework for an optimal use of vehicle actuators is Toyota VDIM (Vehicle Dynamics Integrated Management, [5]), but it let the driver to decide whether to enable or not in very specific situations. Moreover, it does not allow the possibility to communicate with other vehicles (V2V) or with the infrastructure (V2I). In other words, fully automated driving was not considered in none of these cases.

* This work was supported by the Plan Nacional, under the project Tránsito (TRA2008-06602-C03-01), by the Comisión Interministerial de Ciencia y Tecnología under the project GUIADE (Ministerio de Fomento T9/08) and the Ministerio de Ciencia e Innovación under the project CityElec (PS-370000-2009-4).

R. Moreno-Díaz et al. (Eds.): EUROCAST 2011, Part II, LNCS 6928, pp. 390–397, 2012.

This paper presents a modular control architecture for automated vehicles to fill this gap. Figure 1 schematizes an adaptation of the classical robotics paradigm perception-reasoning-action to the automatic driving context. The key issues of this contribution are located at the reasoning part, where a first work division exist at the highest level: planning and control. The planning phase will be in charge of deciding the geometrical path to track and the reference speed for each situation. To achieve such a task, all on-board sensors and V2V/V2I communication based information will be intelligently fused -perception stage- to interpret vehicle dynamics in its environment -road state, weather conditions, traffic. From these inputs, the planner will generate adequate references to the most relevant dynamic variables for vehicle handling. From this point, the

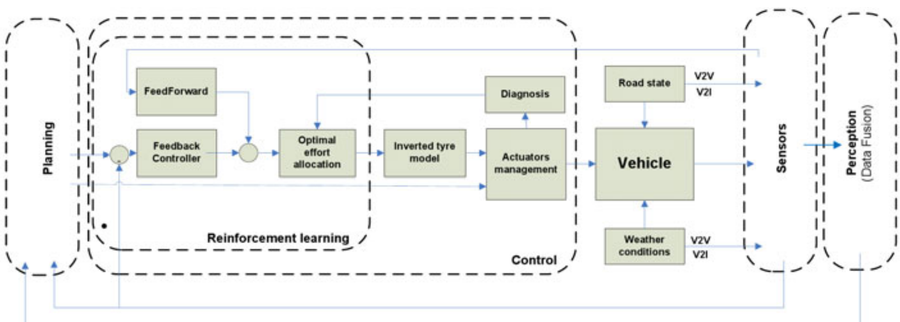

Fig. 1. Modular control architecture for automated vehicles

control layer will decide the best control action for each actuator. This modular control scheme fulfills five fundamental exigences for fully autonomous driving, that are nowadays hot topics for the ITS community and will be detailed in the following sections: efficiency, generality, adaptability, robustness and safety.

The remainder of the article is organized as follows. The basis of the allocation algorithm will be described in 2, where optimality (and therefore efficiency) issues will be discussed. Section 3 will be devoted to briefly highlight the generality of the proposed architecture. Adaptability, robustness, efficiency and safety aspects will be briefly discussed in Sections 4, 5, 6 and 7, respectively. Finally some concluding remarks will be drawn

2 A Modular Architecture: From Chassis to Tyres

The modular control architecture schematized in Fig. 1 can be reinterpreted as a sequence of tasks, as depicted in Fig. 2. Thus, from driver inputs (generally, brake, throttle and steering angle) and taking into account the perceived information from the environment (road state, weather conditions, surrounding vehicles...), a path and trajectory planner quantifies the vehicle targets, and finally the global control scheme comes

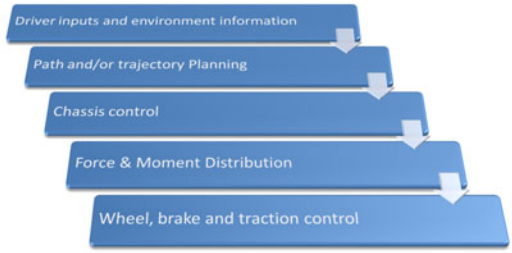

Fig. 2. Flowchart of the proposed control architecture

into the scene. In the latter, a three layer bidirectional flowchart is used to react as precisely and efficiently to the targets defined in the first two layers.

Firstly, an virtual feedback control loop is set to track the three most influent variables in vehicle dynamics: longitudinal velocity V_x, lateral velocity V_y and yaw rate $\dot{\psi}$. Thereafter, the required force and moment in the center of gravity is distributed by an optimal allocation algorithm to the tires. Finally the resulting angles and traction forces for each wheel are controlled by local controllers of each available actuator.

2.1 Virtual Feedback Control for Global Chassis Control

As previously mentioned, this first loop determines the necessary efforts (F_x and F_y) and torque (M_z) at the center of gravity to track the pre-computed V_x, V_y and $\dot{\psi}$ references. To that end, Newton's second law of motion

$$\dot{V}_{x_{ref}} = \frac{1}{M}F_x,\ \dot{V}_{y_{ref}} = \frac{1}{M}F_y,\ \dot{\psi}_{ref} = \frac{1}{I_z}M_z \tag{1}$$

where M and I_z are the mass and the inertia moment of the vehicle. Since equation (1) is to simplistic to accurately describe vehicle's motion, three SISO robust feedback controllers (u_x, u_y, u_ψ) are added to compensate in finite time the divergences between the references and the measured values V_x, V_y and $\dot{\psi}$

$$\begin{aligned}
u_x &= F_x = M\dot{V}_{x_{ref}} + \text{i-PID}(e_x), e_x = V_{x_{ref}} - V_x \\
u_y &= F_x = M\dot{V}_{y_{ref}} + \text{i-PID}(e_y), e_y = V_{y_{ref}} - V_y \\
u_\psi &= M_z = I_z\ddot{\psi}_{ref} + \text{i-PID}(e_y), e_\psi = \dot{\psi}_{ref} - \dot{\psi}
\end{aligned} \tag{2}$$

being i-PID a robust and efficient control technique that is detailed in Section 5.

2.2 Optimal Effort Distribution

According to [3], the best solution to this problem is obtained by collectively minimizing the instantaneous friction coefficient γ_i of all wheels

$$\gamma_i = \frac{\sqrt{F_{x_i}^2 + F_{x_i}^2}}{F_{z_i}}, \ i = 1 \dots 4$$

where F_{z_i} is the vertical load of each wheel. Therefore, the minimizing problem can be expressed by normalizing a uniform $\gamma = \gamma_i, \imath = \dots 4$ with the sum of the requested moment M_0 and efforts F_0

$$\min_{q_i,\gamma} \frac{\gamma}{lF_0 + M_0}$$

where l is the wheelbase and q_i are the angles between the vertical mid-plane of each wheel and the resulting tire-road effort vector at the centre of that wheel.

Since this particular problem did not provide satisfying results for a certain number of situations where a supplementary degree of flexibility was necessary, a different optimization problem has been considered

$$\min_{q_i,\gamma_i} \frac{\sum_i (\gamma_i - \bar{\gamma})^2}{lF_0 + M_0}, \ \bar{\gamma} = \sum_i \gamma_i \tag{3}$$

A simplified SQP algorithm is applied to solve (3) in real time with the aim of finding the optimal values of γ_i and q_i at each wheel, that fulfill the following expressions

$$\begin{aligned} F_{x_i} &= \gamma F_{z_i} \cos(q_i + \theta) \\ F_{y_i} &= \gamma F_{z_i} \sin(q_i + \theta) \end{aligned}$$

which provide the required traction efforts at each wheel F_{x_i}. Finally, the steering angles α_i are obtained via the Brush model [6]

$$\begin{aligned} \alpha &= \tan^{-1}\left(\frac{\kappa_s}{\kappa_\alpha} \frac{-\kappa_i \sin(q_i + \theta)}{1 - \kappa_i \cos(q_i + \theta)}\right) \\ \kappa_i &= \frac{3F_i}{\kappa_s}\left(1 - (1 - \gamma)^{\frac{1}{3}}\right) \end{aligned}$$

where κ_s is the slip stiffness and κ_α the cornering stiffness Figure 3 shows two different situations in a 4 wheel driven and steering car (4WS), where the wheel on each load is different (the magenta dashed circles have the same radius). The vehicle is schematically represented in an aerial view, where the target effort torsor is represented in the gravity center by a circle arc representing M_0 and an arrow of magnitude and orientation given by F_x ana F_y. Note that while in the first case, the vehicle attains the maximum grip margin ($\gamma_i = 1, i = 1 \dots 4$), in the second one the efforts distribution is more comfortable and only a 73% of the tire potential is used.

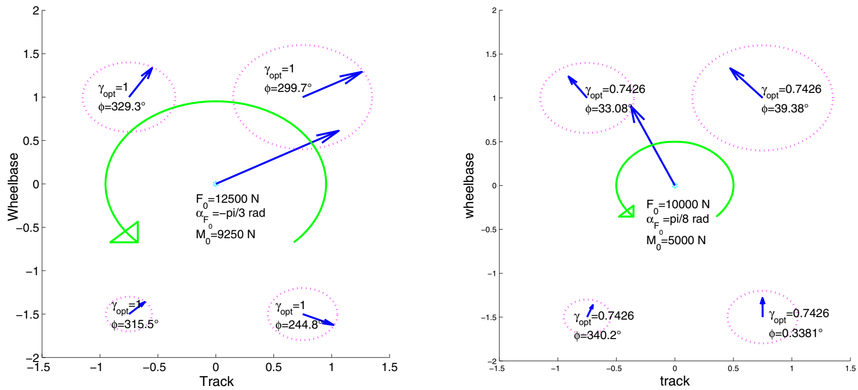

Fig. 3. Examples of efforts allocation with the same load distribution, but with different targets

3 Generality

The proposed control architecture is conceived in such a way that any control technique can be used in both the local and the high level virtual controllers. Moreover, a supplementary feedforward controller can complement the chosen feedback control law (see for instance [8]) The generality is not only related to the architecture modularity, but

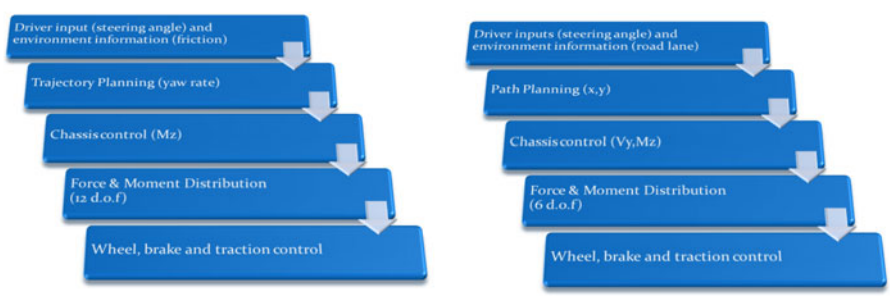

Fig. 4. Flowchart for an ESP in a 4WS vehicle (left) and a Lane Keeping in an AFS car (right)

also to the wide range of application it is useful for. In this connection, Fig. 4 shows how the generic flowchart presented in Fig. 2 may be adapted to two different ADAS with completely different actuator configurations: an ESP-like control for a 4WS vehicle and a Lane Keeping System for a Active Front Steering (AFS) vehicle.

4 Adaptability

Reinforcement Learning [10] permits to adapt on-line the controller parameters to the driver and to the vehicle. As a result, the control layer is able to adapt information provided by the driver (maximum desired velocity, acceleration, suspension behavior...) to each vehicle capabilities at any moment of its lifetime cycle. Figure 5 shows the basis of

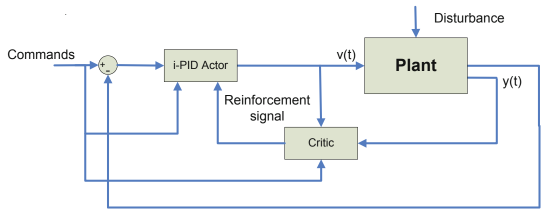

Fig. 5. Reinforcement learning scheme for Global Chassis Control

our on-line learning based control scheme. A critic learns how controller act optimally by intelligently recovering feedbacks from the environment so that desirable actions are reinforced and vice versa. To that end, 12 parameters of virtual control equations(4 control parameters for each one of the 3 degrees of freedom) are continuously tuned to adapt to pre-learnt situations. As a consequence of this, a self learning driver cognition model is transformed into an optimal control problem.

5 Robustness

Tyre and brake wear, pressure, temperature and many other factors decisively influences the dynamic behavior of a vehicle. Such effects are extremely difficult to model, and when a model is available it is not tractable for control based algorithms. Alternatively, robust model based solutions have been tested but they are not always efficient enough, because it is not easy to quantify all the unmodeled dynamics or the parametric uncertainty.

To solve this complex problem, a trajectory tracking based control, and somehow independent of the model, seems an interesting alternative to the aforementioned techniques. Therefore, intelligent PID (i-PID) controllers are used in this work because they combine the well-known PID structure with an "intelligent" term [2] that compensates the effects of nonlinear dynamics, disturbances or uncertain parameters.

As given in [2,9], a finite dimension nonlinear system can be written locally as

$$y^{(\mu)} = F + \alpha u \tag{4}$$

where $\alpha \in \mathbb{R}$ and $\mu \in \mathbb{N}$ are two constant parameters, which do not necessarily represent a physical magnitude, and whose choice is based on the following guidelines

- μ is usually 1 o 2, and it may represent the system order, but not necessarily.
- α should allow F and αu to be of the same order of magnitude.
- The term F is a sort of non-linear black box identifier [2]

If equation (4) is inverted and merged with a PID controller, the resulting i-PID control law yields (see Fig. 6 to clarify how each term is interconnected)

$$u = \frac{1}{\alpha}\left(\dot{V}_s - F\right) + K_P e + K_I \int e + K_D \frac{de}{dt} dt, e = y_r - y \tag{5}$$

where K_P, K_I and $K_D \in \mathbb{R}^+$ are PID tuned gains and y_r is a smoothed reference.

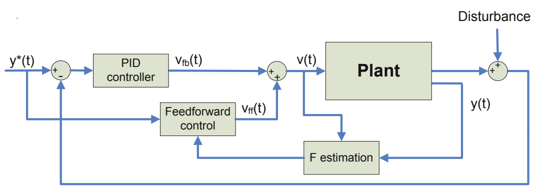

Fig. 6. i-PID control scheme

6 Efficiency

Since the optimization algorithm detailed in Section 2 takes into account the available sensors and actuators, a considerable improvement on energy consumption can be achieved. As a result, this modular architecture is specially interesting for fully actuated electric vehicles, not only in terms of energetic efficiency, but also concerning the maximum resultant forces and moments that can be handled by a vehicle. To reinforce this idea, Ono and coworkers [3] showed that with a very similar architecture the latter could be enhanced in a 18% if the vehicle was 4WS instead of a conventional one.

7 Safety

The presented control allocation scheme, based in the general architecture showed in Fig. 1 facilitates fault tolerant control [7] in case the on board diagnosis tools detects dysfunctional sensors or actuators. Figure 7 shows the basis of a fault tolerant scheme adapted to our GCC architecture, where some critic variables are on-line estimated to evaluate wether residuals are significant or not to adopt a control reconfiguration. If a failure is detected, the effort allocation stage takes into account the dysfunctional componentes to distribute the targets among the properly working actuators.

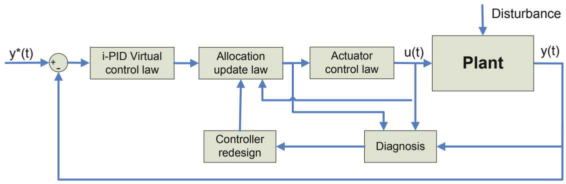

Fig. 7. Fault tolerant control for an effort allocation based control scheme

8 Concluding Remarks

A modular and generic architecture has been proposed to deal with Global Chassis Control. Different techniques are combined to provide generality, robustness, adaptability, efficiency and safety to the control architecture of any kind of wheeled vehicle. The proposed architecture is being evaluated with a realistic simulator (specially the reinforcement learning part). Moreover, the first experimental results will be soon available for very different vehicles.

References

1. Chou, H., D'Andréa-Novel, B.: Global vehicle control using differential braking torques and active suspension forces. Vehicle System Dynamics 43(4), 261–284 (2005)
2. Fliess, M., Join, C.: Intelligent PID Controllers. In: Proc. of 16th Mediterrean Conf. on Control and Automation, Ajaccio, France (2008)
3. Ono, E., Hattoria, Y., Muragishia, Y., Koibuchi, K.: Vehicle dynamics integrated control for four-wheel-distributed steering and four-wheel-distributed traction/braking systems. Vehicle System Dynamics 44(2), 139–151 (2006)
4. Poussot-Vassal, C., Sename, O., Dugard, L.: A LPV/H_∞ Global Chassis Controller for handling improvements involving braking and steering systems. In: IEEE 47th Conference on Decision and Control, pp. 5366–5371 (2008)
5. Suzumura, M., Fukatani, K., Asada, H.: Current State of and Prospects for the Vehicle Dynamics Integrated Management System (VDIM). Toyota Technical Review 55(222) (2007)
6. Svendenius, J., Gäfvert, M.: A semi-empirical dynamic tire model for combined-slip forces. Vehicle System Dynamics 44(2), 189–208 (2006)
7. Tondel, P., Johansen, T.A.: Control allocation for yaw stabilization in automotive vehicles using multiparametric nonlinear programming. In: Proc. of the American Control Conference, June 8-10, pp. 453–458 (2005)
8. Villagra, J., d'Andrea-Novel, B., Mounier, H., Pengov, M.: Flatness-Based Vehicle Steering Control Strategy With SDRE Feedback Gains Tuned Via a Sensitivity Approach. IEEE Transactions on Control Systems Technology 15(3), 554–565 (2007)
9. Villagra, J., Milanes, V., Pérez, J., de Pedro, T.: Control basado en pid inteligentes: Aplicación al control de crucero de un vehículo a bajas velocidades. Revista Iberoamericana de Automática e Informática Industrial 7(4), 44–52 (2010)
10. Wang, X.S., Cheng, Y.H., Sun, W.: A Proposal of Adaptive PID Controller Based on Reinforcement Learning. Journal of China University of Mining and Technology 17(1), 40–44 (2007)