Cooperative guidance of Lego Mindstorms NXT mobile robots

Julien Marzat, Hélène Piet-Lahanier, Arthur Kahn
ONERA – The French Aerospace Lab, F-91123 Palaiseau, France,
{julien.marzat,helene.piet-lahanier,arthur.kahn}@onera.fr

Keywords: autonomous robots, cooperative control, model predictive control, source localization

Abstract: This paper presents experimental results of cooperative guidance laws embedded on Lego Mindstorms NXT mobile robots for two types of missions. The first one is navigation to a waypoint as a fleet with collision and obstacle avoidance, following a model predictive control (MPC) framework. The second one is source localization, i.e., finding the maximum of a potential field, for which a distributed estimation and control strategy is proposed. Experiments show the ability to perform the two missions on these basic mobile robots, in spite of their limited computational resources. In particular, the search for the optimal control sequence through a dedicated discretization of the command space makes it possible to implement real-time MPC.

1 INTRODUCTION

Cooperation between autonomous mobile robots is a challenging task which is currently a very active research field. Several approaches have been suggested for designing decentralized control laws that allow each vehicle to follow a trajectory without affecting the performances of the other vehicles of the fleet, while achieving a required cooperative task [Murray, 2007]. An efficient method is to evaluate distributively a common criterion based on each vehicle action and measurements and the interaction between vehicles. The determination of the control laws can be be derived relying on approaches such as model predictive control (MPC) [Dunbar and Murray, 2006] with individual optimization.

This approach has been validated by simulation but requires experimental testing to ensure that it can be embedded on autonomous platforms. The chosen test system is the low-cost Lego Mindstorms NXT. This popular platform has already been used for basic control of a single robot [Costa et al., 2011, Valera et al., 2011] in particular for control education [Canale and Brunet, 2013] and for testing estimation or localization algorithms [Pinto et al., 2012]. A few attempts at cooperativeness have been made with these robots mainly for coordination and flocking [Benedettelli et al., 2009, Maze et al., 2012]. These works did not consider MPC as a potential solution for fleet coordination and collision avoidance, since it is usually regarded as computationally too expensive. However, our experiments show that it is possible to embed an MPC algorithm on robots with limited computing capacities if a suboptimal search procedure is adopted.

Two classical missions illustrate experimentation of embedded cooperative control and estimation on Lego Mindstorms NXT (see description of the robots in Section 2). The first one (Section 3) involves the navigation to waypoints by a fleet of robots, while avoiding collisions between agents and with obstacles. The second one (Section 4) consists in finding the localization of a source as the maximum of a potential field based on local measurements performed by the agents while moving in fleet.

2 LEGO MINDSTORMS NXT MOBILE ROBOTS

A fleet of $N$ Mindstorms robots has been considered, all built according to a two-wheel differential structure (Figure 1). The dynamical model of the $i$-th vehicle is

\[
\begin{align*}
x_i(k+1) &= x_i(k) + \Delta v_i(k) \cos(\chi_i(k)) \\
y_i(k+1) &= y_i(k) + \Delta v_i(k) \sin(\chi_i(k)) \\
\chi_i(k+1) &= \chi_i(k) + \Delta u_i^\omega(k)
\end{align*}
\]

where $\mathbf{p}_i = [x_i, y_i]^T$ is the vehicle position and $\chi_i$ its direction angle, which form the state vector $\mathbf{x}_i = [x_i, y_i, \chi_i]^T$; $\Delta t$ is the sampling timestep. The velocity $v_i$ was set to a constant value for simplicity, the only control input is thus the rotational speed $u_i = u_i^\omega$, constrained between $\pm \omega_{\text{max}}$. 


For the \( i \)-th robot, practical control of linear and angular velocity via the controllable rotation speeds of the wheels \( \omega_l \) and \( \omega_r \) is achieved by

\[
\begin{align*}
\nu_i &= \frac{(\omega_l + \omega_r)}{2r} \\
\omega^i_l &= \frac{(\omega^i_r - \omega^i_l)}{2L} \\
\end{align*}
\]

(2)

where \( r \) is the wheel radius and \( L \) the half-axis length. Localization of each robot is performed using odometry based upon the previous model (embedded wheel sensors have an accuracy of 1 degree). Since the test missions were short, this localization method was deemed sufficient for estimating the position of the robots in spite of the error accumulated by odometry.

**3 FLEET NAVIGATION WITH COLLISION AND OBSTACLE AVOIDANCE BY MPC**

The first mission considered is the navigation of a fleet of Mindstorms toward a waypoint, with collision and obstacle avoidance. This is tackled using a simplified version of the MPC framework initially defined in [Rocchefort et al., 2012], where more theoretical details can be found.

### 3.1 MPC method

In distributed MPC [Scattolini, 2009], each vehicle computes its control inputs at each timestep as a solution of an optimization problem over the future predicted trajectory. For tractability reasons, finite prediction and control horizon lengths, respectively denoted \( H_p \) and \( H_c \) are used. Future control inputs and resulting state trajectories are

\[
\begin{align*}
U_k &= ((u_1(k)) \tau, (u_1(k + 1)) \tau, ..., (u_1(k + H_r - 1)) \tau)^T \\
X_k &= ((x_1(k + 1)) \tau, (x_1(k + 2)) \tau, ..., (x_1(k + H_p)) \tau)^T \\
\end{align*}
\]

When \( H_c < H_p \), we assume that the control inputs are null after \( H_c \) steps. Once the optimal input sequence \( U^*_k \) has been computed, each vehicle communicates its predicted trajectory to the rest of the fleet and applies the first entry \( u^*_i(k) \). The optimization problem at time \( k \) is stated as

\[
\begin{align*}
\text{minimize } & J_i(U_i, X_i) \\
\text{over } & U_i \subset U_i^{H_c} \\
\text{with } & x_i(t) \text{ satisfying (1), } \forall t \in [k + 1; k + H_p] \\
\end{align*}
\]

(3)

**J** \( i \) is the cost function associated with vehicle \( i \). The constraints coupling the dynamics of the vehicles, such as collision avoidance, are taken into account by penalization. At the next timestep, each vehicle solves its optimization problem considering that the other vehicles follow their predicted trajectories. \( J_i \) comprises a navigation cost \( J_i^{\text{nav}} \), a safety cost \( J_i^{\text{safety}} \) and a control cost \( J_i^c \) such that

\[
J_i(k) = J_i^{\text{nav}}(k) + J_i^{\text{safety}}(k) + J_i^c(k) \\
\]

(4)

Formulation of each cost function is presented in the following subsections. Weighting coefficients \( W^* \) are tuned to set relative priorities between each aspect of the mission.

#### 3.1.1 Navigation cost

The navigation cost \( J_i^{\text{nav}} \) aims at controlling how vehicles navigate to waypoints. It is decomposed into

\[
J_i^{\text{nav}}(k) = J_i^{\text{nav direct}}(k) + J_i^{\text{nav fleet}}(k) \\
\]

(5)

The reference trajectory to the next waypoint \( p_p \) is composed of points \( p^{ref}_i(nk)(n \in [k + 1, k + H_p]) \). They correspond to positions that vehicle \( i \) would reach at timestep \( n \) if moving along a straight line to
$p_o$ at nominal velocity $v_i$. These reference points are thus defined by (6) and the associated cost $J_{i,nav}^{predirect}$ is given by (7). The notation $\hat{p}_i(n|k)$ represents the predicted position of robot $i$ at time $n$, starting from instant $k$.

$$p_{i,p}^{ref}(n|k) = p_i(k) + (n-k)\Delta v_i \frac{p_i(k) - p_o}{\|p_i(k) - p_o\|}$$  \hspace{1cm} (6)

$$J_{i,nav}^{predirect}(k) = W_{nd}^i \sum_{n=k+1}^{k+H_p} \|\hat{p}_i(n|k) - p_{\hat{p}_i}^{ref}(n|k)\|^2$$  \hspace{1cm} (7)

$J_{i,nav,fleet}$ aims at keeping the vehicles together as a fleet. Its definition penalizes the predicted distance $\hat{d}_{ij}(n|k) = \|\hat{p}_i(n|k) - \hat{p}_j(n|k)\|$ between vehicles $i$ and $j$ ($i \neq j$), as

$$J_{i,nav,fleet}(k) = W_{\alpha}^i \sum_{j=1}^{N} \sum_{n=k+1}^{k+H_p} \frac{1 + \tanh\left(\left(\hat{d}_{ij}(n|k) - \beta_{ij}^f\right)\alpha_{ij}^f\right)}{2}$$  \hspace{1cm} (8)

where coefficients $\beta_{ij}^f$ and $\alpha_{ij}^f$ are defined by

$$\beta_{ij}^f = 6(d_{\alpha} - d_{\alpha}^{safe})^{-1}, \quad \alpha_{ij}^f = \frac{1}{2}(d_{\alpha}^{des} + d_{\alpha}^{safe})$$  \hspace{1cm} (9)

The coefficient $d_{\alpha}^{des}$ defines a desired distance between the vehicles inside the fleet whereas $d_{\alpha}^{safe}$ is the maximum distance allowed between vehicles of the fleet. Vehicles $j$ ($i \neq j$) beyond this maximum distance are not considered any more by vehicle $i$. This represents for example limited communication and/or sensing ranges.

### 3.1.2 Safety cost

The safety cost $J_{i,safety}$ aims at avoiding collisions with obstacles, and between vehicles within the fleet. It is made of two costs

$$J_{i,safety}(k) = J_{i,safety,veh}(k) + J_{i,safety,obs}(k)$$  \hspace{1cm} (10)

The first cost deals with collision avoidance between vehicles by penalizing the predicted distance $\hat{d}_{ij}$ between them:

$$J_{i,safety,veh}(k) = W_{\alpha}^{io} \sum_{j=1}^{N} \sum_{n=k+1}^{k+H_p} \frac{1 - \tanh\left(\left(\hat{d}_{ij}(n|k) - \beta_{ij}^v\right)\alpha_{ij}^v\right)}{2}$$  \hspace{1cm} (11)

where

$$\beta_{ij}^v = 6(d_{\alpha}^{des} - d_{\alpha}^{safe})^{-1}, \quad \alpha_{ij}^v = \frac{1}{2}(d_{\alpha}^{des} + d_{\alpha}^{safe})$$  \hspace{1cm} (12)

$J_{i,safety,veh}$ with respect to the distance $\hat{d}_{ij}$ are plotted in Figure 2.

### 3.1.3 Control cost

As traditionally introduced in MPC, the control cost $J_{i,control}$ aims at limiting the control effort, and thus energy consumption of vehicle $i$. It is simply defined as

$$J_{i,control}(k) = W_{\alpha}^{io} \sum_{n=k+1}^{k+H_p} (u_{\alpha}^{i}(n))^2$$  \hspace{1cm} (15)

### 3.1.4 Online computation of control inputs

As the computational cost should be reduced to cope with the robot resources, we limit the search to a finite set $S$ of candidate control sequences. MPC guarantees stability even if the resulting sequences are not optimal but ensures the steady decrease of the cost, as shown in [Scokaert et al., 1999]. At each timestep, the control problem (3) is solved as follows:

1. using a model of the vehicle dynamics, predict the effect of each control sequence of the set of candidates $S$ on the state of the vehicle;
2. compute the cost $J_i$ corresponding to each remaining candidate control sequence;
3. select the control sequence with smallest cost.

The distribution of the candidate control sequences is chosen so as to limit their number while providing a good coverage of the control space, according to the following three rules:
1. the set $S$ of candidates includes the extreme control inputs to exploit the full vehicle potential;
2. the set $S$ of candidates includes the null control input to allow to continue along the same path;
3. candidates are distributed over the entire control space with an increased density around the null control input.

The set of control inputs for the problem at hand reduces to the discretization, with a varying step $\gamma$,

$S = \left\{ \frac{2\pi \gamma}{\eta^w} \right\}$ with $\gamma \in [1, \eta^o]$, \hspace{1cm} (16)

where $\eta^o$ should be chosen to keep the computation of predicted trajectories within the duration of a timestep.

3.2 Experimental results

The chosen discretization strategy to select the optimal control input makes it possible to embed the MPC guidance laws on such mobile robots with limited capabilities, which is otherwise not feasible. The parameters for the experiments presented here were $v_i = 0.1 \text{ m/s}$, $\omega_{\max} = 2.5 \text{ rad/s}$ (consistent with the actual motor limitations), $\Delta t = 0.3 \text{ s}$, $H_c = 4$, $H_p = 8$ and $\eta^o = 11$. The computation time achieved at each iteration was constant and comfortably smaller than $\Delta t$.

Collision avoidance by MPC To validate the communication capabilities of the vehicles and the safety cost ensuring collision avoidance, the first scenario required each robot to go to the initial position of the other robot while avoiding it on the way. The global cost function was thus

$$J_1^i = J_{\text{nav}, \text{direct}}^i + J_{\text{safe}, \text{veh}}^i + J_{\text{obs}}^i$$ \hspace{1cm} (17)

where the navigation cost $J_{\text{nav}, \text{direct}}^i$ for robot $i$ steered it to reach the initial position $p_j$ of robot $j$. The obtained trajectories are given in Figure 3, where it can be observed that the two robots reach their destination with good accuracy and collision avoidance is effective with the desired distance ($d_{\text{des}}^v = 0.5 \text{ m}$ here).

Safe fleet navigation by MPC In addition to the two costs tested in the first experiment, the second scenario involved a fleet behavior (cost $J_{\text{nav, fleet}}^i$) with a desired distance between the vehicles equal to 0.2 m and an obstacle (whose position is known) to avoid thanks to the cost $J_{\text{safe, obs}}^i$. For this experiment, the global cost function was

$$J_2^i = J_{\text{nav}, \text{direct}}^i + J_{\text{nav, fleet}}^i + J_{\text{safe, veh}}^i + J_{\text{safe, obs}}^i + J_{\text{obs}}^i$$ \hspace{1cm} (18)

where now both vehicles were required to go to the same waypoint as a fleet. The obtained trajectories (Figure 4) illustrate the fleet behavior of the vehicles (distance $d_{\text{des}}^v = 0.2 \text{ m}$ is respected) before they encounter the obstacle and avoid it (with a safety distance $d_{\text{safe}}^v = 0.1 \text{ m}$) and finally head toward the same waypoint.

4 SOURCE LOCALIZATION

The second scenario is the localization of the maximum of a potential field $\phi$ using the distributed measurements acquired by the robots of the fleet. This is a problem where cooperative distributed estimation is necessary, since a single robot would have difficulties to localize the source with a single sensor (information would be gathered only on its own path).

The MPC costs obtained for Scenario 1 are still valid for fleet management and collision or obstacle avoidance, but the navigation cost should now be replaced by the control described in Section 4.2.

4.1 Distributed estimation

The field $\phi$, function of position $p_i$, is assumed to be time-invariant and concave (maximum is unique,
second-order derivative \( \nabla^2 \phi \) is negative everywhere. The source-localization algorithm consists in estimating locally the spatial gradient \( \nabla \phi \) of the field \( \phi \) using the fleet of robots, each of the vehicle being able to measure (with identical sensors) the field value at its current position and broadcast it to the rest of the fleet. The fleet of vehicles will then move along the direction of the estimated gradient to head toward the maximum. For estimating the gradient, a vector \( y \) of \( m \) measurements is considered

\[
y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}
\]  

(19)

where \( y_j = \phi(p_j) + e_j \) is the measurement of the field \( \phi \) gathered at a vehicle position \( p_j \) and \( e_j \) is the measurement noise assumed to be distributed according to a zero-mean Gaussian distribution of variance \( \sigma^2_e \). The different measurements \( y_j \) \((j = 1, \ldots, m)\) are acquired by the robots of the fleet and can possibly be measurements at successive time steps. For instance, if \( m \) is chosen to be greater than the number of robots \( N \), a design choice should be made to use past measurements of the robots associated to successive positions (with a timestep large enough to avoid ill-conditioning).

The actual function \( \phi \) is unknown, so a linear estimation model is considered around a given position \( p_0 = [x_0, y_0]^T \) (corresponding to a virtual position within the fleet) as

\[
\phi(p) = \phi(p_0) + (x - x_0) \nabla \phi_{x_0} \nabla \phi_{y_0},
\]

which can be written as

\[
\phi(p) = [1 \ x - x_0 \ y - y_0] \beta,
\]  

(21)

where \( \beta = [\phi(p_0), \nabla \phi_{x_0}, \nabla \phi_{y_0}]^T \) is an unknown vector of parameters to be estimated using the measurement vector \( y \) and the corresponding positions \( p_j \) \((j = 1, \ldots, m)\). The model with the \( m \) measurements can then be written as \( y = H \beta \), where

\[
H = \begin{bmatrix} 1 & x_1 - x_0 & y_1 - y_0 \\ \vdots & \vdots & \vdots \\ 1 & x_m - x_0 & y_m - y_0 \end{bmatrix}
\]  

(22)

The least-square estimate of \( \beta \) is

\[
\hat{\beta} = (H^T H)^{-1} H^T y
\]  

(23)

and the estimated gradient \( \hat{\nabla} \phi = [\hat{\nabla} \phi_{x_0}, \hat{\nabla} \phi_{y_0}]^T \) can be used as the steepest climbing direction to guide the fleet toward the unique maximum of the potential field. Note that with the assumed linear model, the estimated gradient \( \hat{\nabla} \phi \) is independent of the reference point \( p_0 \) where it is computed.

### 4.2 Fleet control for maximum seeking

The fleet should then be guided toward the maximum by following the direction indicated by the estimated gradient. In a distributed scheme, each vehicle estimates the same gradient value using the measurements of all the fleet, then aligns its velocity on this direction. A speed control is considered as \( p_i = \nabla \phi \parallel \nabla \phi \parallel_v v_i \) with \( v_i > 0 \), assuming that the estimation error \( \hat{\nabla} \phi - \nabla \phi \) is small enough. It could then be proven that the robots will converge to the position of the unique maximum \( p^* \) by considering the Lyapunov function

\[
V = (p^* - p_i)^T (p^* - p_i),
\]

whose derivative along the system trajectory is

\[
\dot{V} = - (p^* - p_i)^T \nabla \phi \parallel \nabla \phi \parallel_v v_i. \]

The Taylor expansion of the actual field \( \phi \) at position \( p_i \) evaluated at the maximum position \( p^* \) yields

\[
\phi(p^*) = \phi(p_i) + (p^* - p_i)^T (\nabla \phi + \nabla^2 \phi(\xi)(p^* - p_i)),
\]

(26)

where \( \xi \) belongs to the 2D interval \([p_i, p^*]\). Then

\[
\dot{V} = (p_i - p^*)^T \nabla \phi \parallel \nabla \phi \parallel_v v_i,
\]

\[
= (\phi(p_i) - \phi(p^*)) \parallel \nabla \phi \parallel_v v_i
\]

\[
+ (p_i - p^*)^T \nabla^2 \phi(\xi)(p_i - p^*) \parallel \nabla \phi \parallel_v v_i.
\]

(27)

Since the field is concave, \( \nabla^2 \phi < 0 \) everywhere and since \( p^* \) is the maximum location, \( \phi(p_i) - \phi(p^*) \leq 0 \). \( v_i \) is chosen positive, therefore \( \dot{V} < 0 \) and the gradient climbing control makes the vehicles converge toward the maximum of the field.

### 4.3 Experimental results

A grayscale map is considered as a potential field for the experiments, the maximum being located at the darkest spot. In addition to their positions, the robots now also share their measured gray level (using the Lego color sensor). A fleet of three robots is considered and the mission stops when the measurement of one of the robots reaches a predefined target value (another possible stopping condition could be a threshold on the gradient norm). For estimation purpose, 3 points are sufficient to estimate the gradient, 6 can also be considered for better redundancy by taking into account past positions and measurements of the robots. The NXT brick needs 10 ms to invert a 3x3 matrix and 80 ms to solve the least-square estimation problem (23) with 6 measurements, which is compatible with the chosen timestep of 300 ms. Figure 5 illustrates the gradient ascent in the joint space of positions and gray level (mapped between 0 and 1 from lighter to darker). The fleet successfully finds the location of the maximum.
5 CONCLUSIONS AND PERSPECTIVES

The experiments reported in this paper have highlighted the possibility to control a fleet of Lego Mindstorms NXT to fulfill two types of missions with these autonomous vehicles, notwithstanding their computational capabilities. A first scenario has shown that MPC can be a flexible solution to deal with fleet management and collision avoidance between vehicles and with obstacles. An efficient discretization strategy allows the MPC to find an efficient control sequence within constrained time. In a second scenario, a decentralized estimation and control scheme to find the maximum of a potential field has been presented. It involved linear parameter estimation to obtain the gradient of the field and a gradient-ascent control proven to converge to the actual maximum location. Implementation of the two strategies on the fleet of Lego Mindstorms NXT was successful, which shows the interest of these platforms as a practical testbed for cooperative estimation and control under strict implementation constraints.

ACKNOWLEDGEMENTS

The authors would like to thank Guillaume Broussin and Mathieu Touchard, who contributed to these experiments during their internship at ONERA.

REFERENCES


