Optimization based Multi-Robot Localization with Constraints

(Extended Abstract)

Sen Wang, Ling Chen and Huosheng Hu
School of Computer Science and Electronic Engineering, University of Essex, UK
Email: {swangi, lcheno, hhu}@essex.ac.uk

***Abstract*—It has been known that multi-robot localization is superior to single-robot localization in terms of localization accuracy. However, there is no work clearly examining why and how multi-robot cooperation achieves this in the context of optimization based methods. This paper studies a constrained optimization based multi-robot localization algorithm in the perspective of Fisher Information Matrix to provide some novel insights on why and how multi-robot cooperation and optimization constraints are able to improve localization accuracy.**

I. INTRODUCTION

Multi-robot localization (MRL) is to provide pose estimate for a robot team and, more importantly, improve its accuracy by sharing information [3]. It has been known that MRL is superior to single robot based localization in terms of localization accuracy [1]. However, it is not clear why and how the multi-robot cooperation achieves this in the context of optimization although optimization based methods are very popular to solve the MRL problem.

Many MRL systems have practical limitations (e.g., feasible operating area) on states and disturbances, which can be considered as constraints. However, few MRL algorithms consider these constraints. In the field of estimation, it has been acknowledged that incorporating constraints can improve results [2]. Therefore, it is natural to ask why and how the constraints are able to benefit MRL.

In this paper, these two questions are explicitly answered in the perspective of Fisher Information Matrix (FIM), providing some novel insights on MRL systems.

II. METHOD

In a MRL system, N robots and a single mobile beacon cooperate with each other. They can measure ranges between each other. The neighborhood set of robot i at time t is represented as $\mathbf{N}_{i_t}$ with size n_{i_t}. The process model of the ith robot at time t is

$$\mathbf{x}_{i_{t+1}} = f(\mathbf{x}_{i_t}, \hat{\mathbf{u}}_{i_t}, \mathbf{w}_{i_t})$$

where $\hat{\mathbf{u}}_{i_t}$ is robot motion, $\mathbf{w}_{i_t} \sim \mathcal{N}(0, \mathbf{Q}^c_{i_t})$ is an additive Gaussian noise, $\mathbf{F}_{i_t}$ and $\mathbf{G}_{i_t}$ are the Jacobian matrices with respect to $\mathbf{x}_{i_t}$ and $\mathbf{w}_{i_t}$, respectively. The range measurement of robot i to robot q (or beacon $\mathbf{x}_{b_t}$) is

$$\mathbf{z}^{iq}_{r,t} = h_r(\mathbf{x}_{i_t}, \mathbf{x}_{q_t}) + \mathbf{v}^{iq}_{r,t}, \quad q \in \mathbf{N}_{i_t}$$

where $\mathbf{v}^{iq}_{r,t} \sim \mathcal{N}(0, \mathbf{R}^{iq}_{r,t})$. We denote ${}^{i}\mathbf{H}^{iq}_{r,t}$ and ${}^{q}\mathbf{H}^{iq}_{r,t}$ as the Jacobian matrices with respect to $\mathbf{x}_{i_t}$ and $\mathbf{x}_{q_t}$, respectively.

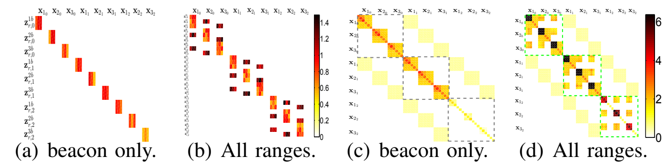


Fig. 1. $\mathbf{H}$ and $\mathbf{D}$ in different situations at step 2. (a) and (b) are $\mathbf{H}$, while (c) and (d) are $\mathbf{D}$. The dashed square indicates a time slot.

The Maximum a Posteriori (MAP) based MRL algorithm with constraints is

$$\begin{aligned} \underset{\{\mathbf{x}_t\}_0^k}{\text{minimize}} \quad & \boldsymbol{\Phi}_r(\{\mathbf{x}_t\}_0^k) + \boldsymbol{\Phi}_p(\{\mathbf{x}_t\}_0^k) + \boldsymbol{\Phi}_0(\{\mathbf{x}_t\}_0^k) \\ \text{subject to} \quad & g_l(\mathbf{x}, \mathbf{u}, \mathbf{w}) \preceq \mathbf{d}_l, \quad l = 0, \cdots, m \end{aligned} \tag{1}$$

where

$$\boldsymbol{\Phi}_r(\{\mathbf{x}_t\}_0^k) = \sum_{t=0}^{k}\sum_{i=1}^{N}\sum_{q\in\mathbf{N}_{i_t}} \|\mathbf{z}^{iq}_{r,t} - h_r(\mathbf{x}_{i_t}, \mathbf{x}_{q_t})\|^2_{\mathbf{R}^{iq}_{r,t}}$$

$$\boldsymbol{\Phi}_p(\{\mathbf{x}_t\}_0^k) = \sum_{t=0}^{k-1}\sum_{i=1}^{N} \|\mathbf{x}_{i_{t+1}} - f(\mathbf{x}_{i_t}, \hat{\mathbf{u}}_{i_t})\|^2_{\mathbf{Q}_{i_t}}$$

$$\boldsymbol{\Phi}_0(\{\mathbf{x}_t\}_0^k) = \sum_{i=1}^{N} \|\mathbf{x}_{i_0} - \hat{\mathbf{x}}_{i_0}\|^2_{\boldsymbol{\Pi}_{i_0}}$$

$\mathbf{x}_t = \left[\mathbf{x}^T_{1_t}, \ldots, \mathbf{x}^T_{i_t}, \ldots, \mathbf{x}^T_{N_t}\right]^T$, $\|p\|^2_A = p^T A^{-1} p$, $\mathbf{Q}_{i_t} = \mathbf{G}_{i_t}\mathbf{Q}^c_{i_t}\mathbf{G}^T_{i_t}$, $\mathcal{N}(\hat{\mathbf{x}}_{i_0}, \boldsymbol{\Pi}_{i_0})$ is prior knowledge on the initial state of robot i, $\mathbf{z}_{r,i}$ is all the range measurements at time t, and $\{\cdot\}_a^b$ denotes a set of quantities from time a to b.

When the optimization problem (1) is iteratively solved, in each iteration we have the following problem after linearization:

$$\begin{aligned} \underset{\{\Delta\mathbf{x}_t\}_0^k}{\text{minimize}} \quad & \boldsymbol{\Phi}(\{\Delta\mathbf{x}_t\}_0^k) \\ \text{subject to} \quad & \mathsf{G}\Delta\mathbf{x}_{0:k} \preceq \mathsf{d} \end{aligned} \tag{2}$$

where

$$\boldsymbol{\Phi}(\{\Delta\mathbf{x}_t\}_0^k) = \sum_{t=0}^{k}\sum_{i=1}^{N}\sum_{q\in\mathbf{N}_{i_t}} \|\mathbf{c}^{iq}_{r,t} - {}^{i}\mathbf{H}^{iq}_{r,t}\Delta\mathbf{x}_{i_t} - {}^{q}\mathbf{H}^{iq}_{r,t}\Delta\mathbf{x}_{q_t}\|^2_{\mathbf{R}^{iq}_{r,t}}$$
$$+ \sum_{t=0}^{k-1}\sum_{i=1}^{N} \|\Delta\mathbf{x}_{i_{t+1}} - \mathbf{F}_{i_t}\Delta\mathbf{x}_{i_t} - \mathbf{a}_{i_{t+1}}\|^2_{\mathbf{Q}_{i_t}} + \sum_{i=1}^{N} \|\Delta\mathbf{x}_{i_0} - \mathbf{a}_{i_0}\|^2_{\boldsymbol{\Pi}_{i_0}},$$

and

$$\mathsf{G} = \begin{bmatrix} \nabla g_0^T \\ \vdots \\ \nabla g_m^T \end{bmatrix}, \qquad \mathsf{d} = \begin{bmatrix} \mathbf{d}_0 - g_0(\mathbf{x}, \mathbf{u}, \mathbf{w}) \\ \vdots \\ \mathbf{d}_m - g_m(\mathbf{x}, \mathbf{u}, \mathbf{w}) \end{bmatrix}$$

where

$\mathbf{c}_{r,t}^{iq} = \mathbf{z}_{r,t}^{iq} - h_r(\mathbf{x}_{i_t}, \mathbf{x}_{q_t}), \mathbf{a}_{i_{t+1}} = f(\mathbf{x}_{i_t}, \hat{\mathbf{u}}_{i_t}) - \mathbf{x}_{i_{t+1}}, \mathbf{a}_{i_0} = \hat{\mathbf{x}}_{i_0} - \mathbf{x}_{i_0}$

Then, $\boldsymbol{\Phi}(\{\Delta\mathbf{x}_t\}_0^k)$ can be rewritten as

$$\boldsymbol{\Phi}(\{\Delta\mathbf{x}_t\}_0^k) = \|\mathbf{C}(\mathbf{A}\Delta\mathbf{x}_{0:k} - \mathbf{b})\|_2^2$$

where

$$\begin{aligned}
&\mathbf{C} = \begin{bmatrix} \mathbf{Q}^{-\frac{1}{2}} & \\ & \mathbf{R}^{-\frac{1}{2}} \end{bmatrix}, \ \mathbf{A} = \begin{bmatrix} \mathbf{F} \\ \mathbf{H} \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} \mathbf{a} \\ \mathbf{c} \end{bmatrix} \\
&\mathbf{Q}^{-\frac{1}{2}} = \mathbf{diag}\left(\boldsymbol{\Pi}_{1_0}^{-\frac{1}{2}}, \cdots, \boldsymbol{\Pi}_{N_0}^{-\frac{1}{2}}, \mathbf{Q}_{1_0}^{-\frac{1}{2}}, \cdots, \mathbf{Q}_{N_{k-1}}^{-\frac{1}{2}}\right) \\
&\mathbf{R}^{-\frac{1}{2}} = \mathbf{diag}\left(\mathbf{R}_{r,0}^{1b\,-\frac{1}{2}}, \mathbf{R}_{r,0}^{12\,-\frac{1}{2}}, \cdots, \mathbf{R}_{r,k}^{(N-1)N\,-\frac{1}{2}}\right) \\
&\mathbf{F} = \mathbf{I} + \begin{bmatrix} \mathbf{0}_{(6\cdot N)\times(6\cdot N\cdot k)} & \mathbf{0}_{(6\cdot N)\times(6\cdot N)} \\ \mathbf{diag}\left(-\mathbf{F}_{1_0}, \cdots, -\mathbf{F}_{N_{k-1}}\right) & \mathbf{0}_{(6\cdot N\cdot k)\times(6\cdot N)} \end{bmatrix} \\
&\mathbf{H} = \mathbf{diag}\left(\mathbf{H}_0, \cdots, \mathbf{H}_t, \cdots, \mathbf{H}_k\right) \\
&\mathbf{H}_t = \begin{bmatrix} {}^1\mathbf{H}_{r,t}^{1b} & & & & \\ {}^1\mathbf{H}_{r,t}^{12} & {}^2\mathbf{H}_{r,t}^{12} & & & \\ \cdots & \cdots & {}^i\mathbf{H}_{r,t}^{iq} & \cdots & {}^q\mathbf{H}_{r,t}^{iq} \\ & & \vdots & & \vdots \end{bmatrix}_{(\sum_{i=1}^N \mathbf{n}_{i_t})\times(6\cdot N)} \\
&\mathbf{a} = \begin{bmatrix} \mathbf{a}_0^T & \cdots & \mathbf{a}_k^T \end{bmatrix}^T \ \mathbf{a}_t = \begin{bmatrix} \mathbf{a}_{1_t}^T & \mathbf{a}_{2_t}^T & \cdots & \mathbf{a}_{N_t}^T \end{bmatrix}^T \\
&\mathbf{c} = \begin{bmatrix} \mathbf{c}_0^T & \cdots & \mathbf{c}_k^T \end{bmatrix}^T \ \mathbf{c}_t = \begin{bmatrix} \mathbf{c}_{r,t}^{1b\,T} & \mathbf{c}_{r,t}^{12\,T} & \cdots & \mathbf{c}_{r,t}^{jq\,T} & \cdots \end{bmatrix}^T
\end{aligned}$$

$\mathbf{diag}(\cdot)$ is a function to map vectors to a block diagonal matrix. The Lagrangian of (2) is

$$\boldsymbol{\mathcal{L}}(\Delta\mathbf{x}_{0:k}, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{\Phi}(\{\Delta\mathbf{x}_t\}_0^k) + \boldsymbol{\lambda}^T(\mathbf{G}\Delta\mathbf{x}_{0:k} - \mathbf{d})$$

where $\boldsymbol{\lambda} \in \mathbb{R}^K, K = \sum_{l=0}^m k_l,$. Then, the KKT conditions of the optimization problem (2) are given by

$$\begin{aligned}
\nabla_{\Delta\mathbf{x}_{0:k}}\boldsymbol{\mathcal{L}} &= (\mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\Delta\mathbf{x}_{0:k} \\
&\quad - (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{c} + \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{a}) + \mathbf{G}^T\boldsymbol{\lambda} = \mathbf{0} \qquad (3) \\
\mathbf{diag}(\boldsymbol{\lambda})&(\mathbf{G}\Delta\mathbf{x}_{0:k} - \mathbf{d}) = \mathbf{0}
\end{aligned}$$

Therefore, Newton step for solving (3) is given by

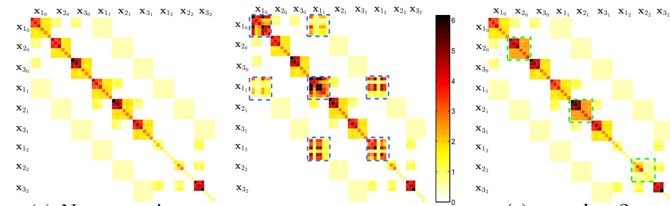
$$\begin{aligned}
&\begin{bmatrix} (\mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) & \mathbf{G}^T \\ \mathbf{diag}(\boldsymbol{\lambda})\mathbf{G} & -\mathbf{diag}(\mathbf{d} - \mathbf{G}\Delta\mathbf{x}_{0:k}) \end{bmatrix} \begin{bmatrix} \delta\mathbf{x}_{0:k} \\ \delta\boldsymbol{\lambda} \end{bmatrix} \\
&= \begin{bmatrix} -(\mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\Delta\mathbf{x}_{0:k} + (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{c} + \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{a}) - \mathbf{G}^T\boldsymbol{\lambda} \\ -\mathbf{diag}(\boldsymbol{\lambda})(\mathbf{G}\Delta\mathbf{x}_{0:k} - \mathbf{d}) \end{bmatrix} \qquad (4)
\end{aligned}$$

Applying Schur Complement to (4), the coefficient of $\delta\mathbf{x}_{0:k}$ can be derived:

$$\begin{aligned}
\mathbf{D} &= \underbrace{(\mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})}_{\text{FIM of unconstrained problem}} + \underbrace{\mathbf{G}^T(\mathbf{diag}(\mathbf{d} - \mathbf{G}\Delta\mathbf{x}_{0:k}))^{-1}\mathbf{diag}(\boldsymbol{\lambda})\mathbf{G}}_{\text{extra part}} \\
&= \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} \\
&\quad + \begin{bmatrix} \mathbf{H} \\ \mathbf{G} \end{bmatrix}^T \begin{bmatrix} \mathbf{R}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{diag}(\mathbf{d} - \mathbf{G}\Delta\mathbf{x}_{0:k}))^{-1}\mathbf{diag}(\boldsymbol{\lambda}) \end{bmatrix} \begin{bmatrix} \mathbf{H} \\ \mathbf{G} \end{bmatrix} \qquad (5)
\end{aligned}$$

III. RESULTS

By examining (5) in the perspective of FIM, $\mathbf{D}$ includes the FIM of the unconstrained optimization problem [4] and an extra part coming from the constraints. In fact, $\mathbf{H}$ represents the structure of the multi-robot cooperation, while the extra part is related with the constraints.



(a) No constraints. (b) on robot 1. (c) on robot 2.

Fig. 2. $\mathbf{D}$ with different constraints.

1) Multi-Robot Cooperation: Assume there are two different situations: each robot can measure ranges to (a) the beacon only; (b) all robots and the beacon. $\mathbf{H}$ and $\mathbf{D}$ of both situations at step 2 are shown in Fig.1. By concentrating on the differences between Fig.1(c)-Fig.1(d), it can be seen that the information on diagonal blocks increases. This means that as more ranges between robots are used, the information on robot positions increases. Moreover, there are some off-diagonal blocks (e.g., cov($\mathbf{x}_{1_0}, \mathbf{x}_{2_0}$)) in $\mathbf{D}$ that appear or become denser, i.e., correlation between robots is introduced or increased. The significance of the information on the diagonal is that the covariance can be reduced in the presence of the off-diagonal elements. Therefore, the localization accuracy can be improved by sharing information between multiple robots.

2) Constraints: Studying (5) can also answer why and how the constraints are useful for MRL algorithms. Fig.2 shows $\mathbf{D}$ with various constraints, i.e., different extra parts in (5). The scenario where no constraint is imposed is a standard of reference, see Fig.2(a). When there are constraints on the velocity of robot 1, the $\mathbf{D}$ in Fig.2(b) shows that the states of robot 1 in different time slots can be constrained, e.g., cov($\mathbf{x}_{1_0}, \mathbf{x}_{1_1}$) increases. This contributed off-diagonal elements can reduce the uncertainties as the multi-robot cooperation does. Moreover, Fig.2(c) is produced with the constraints on the states of robot 2 only. The changes from Fig.2(a) to Fig.2(c) show that the information (uncertainty) on state estimate of robot 2 is improved (decreased). This constraint, for instance, can be used when there is prior knowledge on feasible operating field of robot positions. In fact, the extra part in (5) indicates that the constraints can be conceived as supplementary "sensors", which supply additional information.

REFERENCES

[1] A I Mourikis and S I Roumeliotis. Performance analysis of multirobot cooperative localization. *IEEE Trans. Robot.*, 22(4):666–681, 2006.

[2] Douglas G Robertson and Jay H Lee. On the use of constraints in least squares estimation and control. *Automatica*, 38(7):1113–1123, 2002.

[3] S I Roumeliotis and G A Bekey. Distributed multirobot localization. *IEEE Trans. on Robotic. Autom.*, 18(5):781–795, 2002.

[4] Gabe Sibley, Larry Matthies, and Gaurav Sukhatme. Sliding window filter with application to planetary landing. *J. Field Robot.*, 27(5):587–608, 2010.