How far is SLAM from a linear least squares problem?

Shoudong Huang, Yingwu Lai, Udo Frese and Gamini Dissanayake

***Abstract*— Most people believe SLAM is a complex nonlinear estimation/optimization problem. However, recent research shows that some simple iterative methods based on linearization can sometimes provide surprisingly good solutions to SLAM without being trapped into a local minimum. This demonstrates that hidden structure exists in the SLAM problem that is yet to be understood. In this paper, we first analyze how far SLAM is from a convex optimization problem. Then we show that by properly choosing the state vector, SLAM problem can be formulated as a nonlinear least squares problem with many quadratic terms in the objective function, thus it is clearer how far SLAM is from a linear least squares problem. Furthermore, we explain that how the map joining approaches reduce the nonlinearity/nonconvexity of the SLAM problem.**

I. Introduction

Simultaneous Localization and Mapping (SLAM) has been investigated by robotic researchers for more than 10 years [1]. Although many SLAM algorithms have been developed, most of them treated SLAM as a high dimensional nonlinear estimation/optimization problem. The sparseness of the information matrix in different SLAM formulations is now well understood and exploited thoroughly (e.g. [2][3][4]), but the underlying structure of nonlinearity has not been fully understood yet.

For point feature based SLAM problem, our initial investigation has shown some interesting phenomenon when a simple Gauss-Newton algorithm is applied to solve the SLAM as an optimization problem. For both the Victoria Park data set [5] and the DLR-Spatial-Cognition data set [6], the algorithm can converge with very poor initial values[1]. However, these "magic" convergence happen when the covariances of observations and odometries are set to be identity matrices but not for the original covariance matrices. See Table I for details. Of course, the solution of using identity covariance matrix and that using the original covariance are (slightly) different, as shown in Figure 1.

It is well known that a high dimensional nonlinear optimization problem can have a lot of local minima and a good initial value is critical for an optimization algorithm to converge to the correct solution. On the other hand, linear least squares problems have quadratic objective functions and can be solved in one step without the need of a good initial value. Thus, the above phenomenon shows that SLAM is a very special nonlinear optimization problem that is close to a linear least squares problem in some way.

TABLE I

Convergence of Gauss-Newton algorithm for SLAM with different initial states

data set	covariance	odometry	zero	random
DLR	changed to identity matrix	Yes	Yes	No
DLR	original	Yes	No	No
VicPark	changed to identity matrix	Yes	Yes	Yes
VicPark	original	No	No	No

odometry: initial states from odometry/first observations
zero: initial states are all zeros
random: initial states are randomly given
For Victoria Park data set, the state contains 6898 poses and 299 features
For DLR data set, the state contains 3297 poses and 539 features
'Yes' means the algorithm converges to the correct solution
'No' means the algorithm does not converge to the correct solution

This paper tries to explain how far SLAM is from a linear least squares problem. We first perform some analysis on the convexity of SLAM problem. Then we show that by using the relative information as state vector, the quadratic part and non-quadratic part are clearly distinguished in the objective function. Moreover, using map joining, the quadratic part is more significant as compared with the non-quadratic part.

The paper is organized as follows. Section II provides some notations used in this paper and states the feature based full SLAM problem. Section III explains the details of the traditional least squares SLAM formulation. In Section IV, the convexity of the traditional least squares SLAM is analyzed. Section V proposes the new least squares SLAM formulation using relative state vector. In Section VI, the advantages of using map joining strategy is explained. Section VII discusses the related work. Finally, Section VIII concludes the paper.

II. Preliminaries

In this paper, different coordinate frames need to be clearly distinguished. So some special notations are used.

A. Notations

Suppose there is a sequence of 2D robot poses $r_0, r_1, r_2, \cdots$ and a number of 2D point features $f_1, f_2, \cdots$ in the environments. Normally the first robot pose (pose r_0) is chosen as the origin of the global coordinate frame.

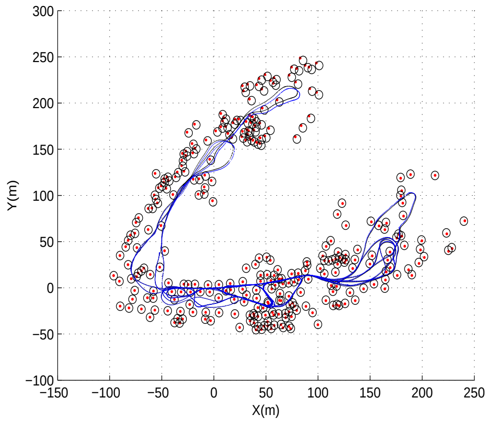
This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government. This work has also been supported by DFG grant SFB/TR 8 Spatial Cognition.

Shoudong Huang and Gamini Dissanayake are with Faculty of Engineering and Information Technology, University of Technology Sydney Australia. {sdhuang,gdissa}@eng.uts.edu.au

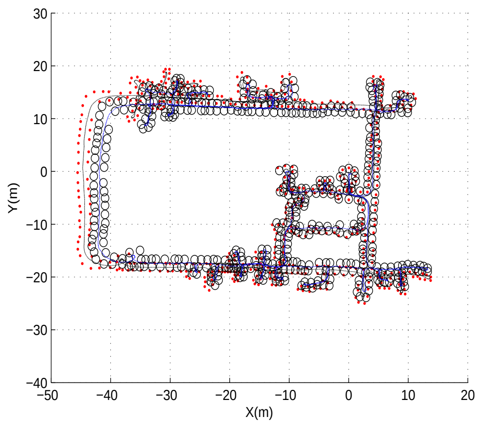
Yingwu Lai is with CNPC Liaoyang Petrochemical Company, Liaoning, China. laiyingwu1964@gmail.com

Udo Frese is with German Research Center for Artificial Intelligence (DFKI) in Bremen, Germany. udo.frese@dfki.de

[1]Here the preprocessed data is used where data association is provided. The preprocessed data is available on OpenSLAM website: http://openslam.org/ under project 2D I-SLSJF.



(a) The Victoria Park data set: 6898 vehicle poses and 299 feature positions. Circle – feature estimate using identity covariance, dot – feature estimate using original covariance



(b) The DLR data set: 3297 robot poses and 539 feature positions. Circle – feature estimate using identity covariance, dot – feature estimate using original covariance

Fig. 1. The least squares SLAM results with original and identity covariance matrix

The following notations are used in this paper to make clear what coordinate frames are used.

$X_{f_j}^{r_i} = (x_{f_j}^{r_i}, y_{f_j}^{r_i})^T$ – the x, y position of feature f_j in the (coordinate) frame defined by pose r_i.

$X_{r_j}^{r_i} = (x_{r_j}^{r_i}, y_{r_j}^{r_i})^T$ – the x, y position of robot pose r_j in the (coordinate) frame defined by pose r_i.

$\phi_{r_j}^{r_i}$ – the orientation of pose r_j in the (coordinate) frame defined by pose r_i.

$R_{r_j}^{r_i}$ – the rotation matrix of the pose r_j in the (coordinate) frame defined by pose r_i.

Note that $X_{f_j}^{r_i}$ and $X_{r_j}^{r_i}$ are both two dimensional vectors. $\phi_{r_j}^{r_i}$ is a scalar and $R_{r_j}^{r_i}$ is a two by two orthogonal matrix.

Some basic equations describing the relationship among the above variables are given below.

For any i, j, k,

$$X_{f_k}^{r_i} = X_{r_j}^{r_i} + R_{r_j}^{r_i} X_{f_k}^{r_j}. \tag{1}$$

$$X_{r_k}^{r_i} = X_{r_j}^{r_i} + R_{r_j}^{r_i} X_{r_k}^{r_j}. \tag{2}$$

$$R_{r_k}^{r_i} = R_{r_j}^{r_i} R_{r_k}^{r_j}. \tag{3}$$

$$\phi_{r_k}^{r_i} = \phi_{r_j}^{r_i} + \phi_{r_k}^{r_j}. \tag{4}$$

$$R_{r_j}^{r_i} = R(\phi_{r_j}^{r_i}) = \begin{bmatrix} \cos\phi_{r_j}^{r_i} & -\sin\phi_{r_j}^{r_i} \\ \sin\phi_{r_j}^{r_i} & \cos\phi_{r_j}^{r_i} \end{bmatrix}. \tag{5}$$

B. Point feature based SLAM problem

Suppose there are N point features $f_1, \cdots, f_N$ that are observed from a sequence of $p+1$ robot poses $r_0, r_1, \cdots, r_p$ with the total number of observations m. Figure 2 shows an example of this scenario with $N = 3$, $p = 4$, and $m = 5$.

In SLAM, there are two kinds of information. Odometry information is the relative pose between two consecutive poses. Observation information is the relative position of the observed feature with respect to the pose where the observation is made.

In this paper, we use Z_j^i to denote the observation made from pose r_i to feature f_j. We use O_j^{j-1} $(1 \le j \le p)$ to denote the odometry information between pose r_{j-1} and pose r_j, $P_{Z_j^i}$ and P_{Oj} are the corresponding covariance matrices of the observation and odometry noises. Here the noises are assumed to be zero-mean Gaussian.

In Figure 2, there are 4 odometries and 5 observations. Using the notations in Section II-A, odometries are the measurements of

$$(x_{r_1}^{r_0}, y_{r_1}^{r_0}, \phi_{r_1}^{r_0}), (x_{r_2}^{r_1}, y_{r_2}^{r_1}, \phi_{r_2}^{r_1}), (x_{r_3}^{r_2}, y_{r_3}^{r_2}, \phi_{r_3}^{r_2}), (x_{r_4}^{r_3}, y_{r_4}^{r_3}, \phi_{r_4}^{r_3}) \tag{6}$$

Observations are the measurements of:

$$(x_{f_1}^{r_0}, y_{f_1}^{r_0}), (x_{f_2}^{r_1}, y_{f_2}^{r_1}), (x_{f_2}^{r_2}, y_{f_2}^{r_2}), (x_{f_3}^{r_3}, y_{f_3}^{r_3}), (x_{f_1}^{r_4}, y_{f_1}^{r_4}) \tag{7}$$

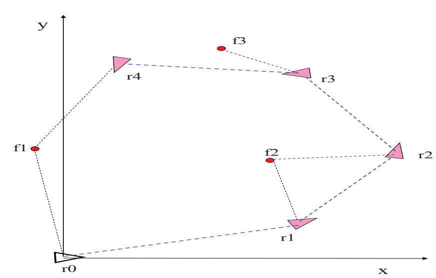


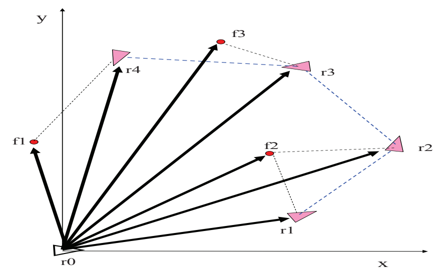
Fig. 2. The SLAM problem with 5 poses, 3 features and 5 observations

The full least squares SLAM formulation [3] is to use the odomety and observation information to estimate all the robot poses and all the feature positions.

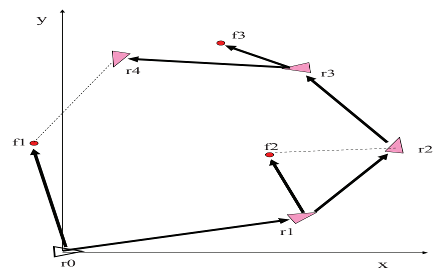
III. Traditional Least Squares SLAM

A. State vector

Traditional least squares SLAM uses the robot poses and the feature positions with respect to robot pose r_0 as the state



(a) The absolute state vector used in traditional least squares SLAM



(b) The relative state vector proposed in this paper

Fig. 3. Different state vectors can be used in SLAM

vector. Using the notations in Section II-A, the state vector is [2]

$$X = (X^{r_0}_{r_1}, \phi^{r_0}_{r_1}, \cdots, X^{r_0}_{r_p}, \phi^{r_0}_{r_p}, X^{r_0}_{f_1}, \cdots, X^{r_0}_{f_N}). \tag{8}$$

An example of the state vector is illustrated in Figure 3(a).

B. Least squares formulation

The full SLAM formulation is to minimize [3]

$$\begin{aligned} F(X) = \;& \sum_{j=1}^{p} (O^{j-1}_j - H^{Oj}(X))^T P^{-1}_{Oj} (O^{j-1}_j - H^{Oj}(X)) \\ & + \sum_{i,j} (Z^i_j - H^{Z^i_j}(X))^T P^{-1}_{Z^i_j} (Z^i_j - H^{Z^i_j}(X)) \end{aligned} \tag{9}$$

where the state variable X is given in (8), O^{j-1}_j $(1 \leq j \leq p)$ are odometries, Z^i_j are observations, and P_{Oj} and $P_{Z^i_j}$ are the corresponding covariance matrices.

In the above least squares SLAM formulation, $H^{Z^i_j}(X)$ and $H^{Oj}(X)$ are the corresponding functions relating Z^i_j and O^{j-1}_j to the state X, most of them are nonlinear functions.

[2]To simplify the notation, sometimes the transpose is omitted.

C. Odometry information function $H^{Oj}(X)$

From the basic equations (2), (4), and (5), the odometry information function is a function of two poses $(X^{r_0}_{r_{j-1}}, \phi^{r_0}_{r_{j-1}})$ and $(X^{r_0}_{r_j}, \phi^{r_0}_{r_j})$ and is given by

$$\begin{aligned} H^{Oj}(X) &= \begin{bmatrix} X^{r_{j-1}}_{r_j} \\ \phi^{r_{j-1}}_{r_j} \end{bmatrix} \\ &= \begin{bmatrix} (R(\phi^{r_0}_{r_{j-1}}))^T (X^{r_0}_{r_j} - X^{r_0}_{r_{j-1}}) \\ \phi^{r_0}_{r_j} - \phi^{r_0}_{r_{j-1}} \end{bmatrix}. \end{aligned} \tag{10}$$

D. Observation information function $H^{Z^i_j}(X)$

The observation is a function of one pose $(X^{r_0}_{r_i}, \phi^{r_0}_{r_i})$ and one feature position $X^{r_0}_{f_j}$ and is given by

$$H^{Z^i_j}(X) = X^{r_i}_{f_j} = (R(\phi^{r_0}_{r_i}))^T (X^{r_0}_{f_j} - X^{r_0}_{r_i}). \tag{11}$$

IV. HOW FAR IS SLAM FROM BEING CONVEX?

A. Definition of convex function

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex, if for any $x, y \in \mathbb{R}^n$,

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y), \;\; \forall \lambda \in (0,1). \tag{12}$$

It is important that a convex function cannot have local minima, since otherwise the connection between two of them would be even smaller by (12).

The surprising convergence result in Table I motivates us to investigate the convexity of the SLAM problem. Figure 4 illustrates the function $F(\lambda X + (1-\lambda)Y)$ in (9) for five random pairs of states X, Y with λ ranging from 0 to 1. The figures indicate that the function $F(X)$ in (9) is not far from being convex for both the Victoria Park data set and DLR data set.

B. Convexity analysis of a single feature observation

We consider a single feature observation Z^j_i and the corresponding term in (9). We assume $P_{Z^i_j} = I$ and define

$$\delta = X^{r_0}_{f_j} - X^{r_0}_{r_i} \tag{13}$$

and omit the indices for brevity. From (11), one term of the objective function related to observation Z is

$$\begin{aligned} f(\phi, \delta) &= [Z - R(\phi)^T \delta]^T [Z - R(\phi)^T \delta] && (14) \\ &= |Z - R(\phi)^T \delta|^2 && (15) \\ &= |R(\phi) Z - \delta|^2 && (16) \end{aligned}$$

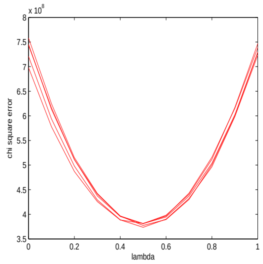
This form is remarkable, because unlike the original form it contains no product of state variables and the only nonlinearity comes from the sines and cosines in $R(\phi)$. This simplification works only for spherical covariance, which explains, why these apparently help convergence in Table I and TORO [7].

Now we denote

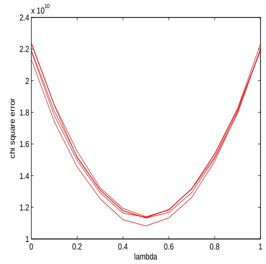
$$\phi = \phi_0 + \tilde{\phi} \tag{17}$$

where ϕ_0 is the estimated value of ϕ (e.g. obtained from odometry) and $\tilde{\phi}$ is the error on the estimation. Then we have

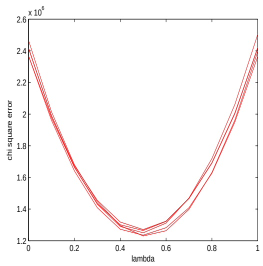
$$f(\tilde{\phi}, \delta) = |R(\tilde{\phi}) R(\phi_0) Z - \delta|^2 = |R(\tilde{\phi}) \hat{Z} - \delta|^2 \tag{18}$$



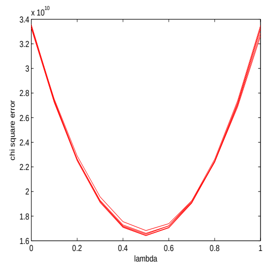
(a) The Victoria Park data with identity covariance matrix (state dimension 21292)



(b) The Victoria Park data with original covariance matrix (state dimension 21292)



(c) The DLR data with identity covariance matrix (state dimension 10969)



(d) The DLR data with original covariance matrix (state dimension 10969)

Fig. 4. Near convexity of the SLAM problem: $F(\lambda X + (1-\lambda)Y)$ in (9) for five random pairs of states X, Y with λ ranging from 0 to 1.

with $\hat{Z} = R(\phi_0)Z$. Note here that $\hat{Z}$ is the relative position (approximately) transferred into the coordinate system defined by r_0.

Further denote $\hat{Z} = [z_x, z_y]^T$, $\delta = [\delta_x, \delta_y]^T$, then we get a function of three variables

$$\begin{aligned} f(\tilde{\phi}, \delta_x, \delta_y) &= |R(\tilde{\phi})\hat{Z} - \delta|^2 && (19)\\ &= (z_x \cos\tilde{\phi} - z_y \sin\tilde{\phi} - \delta_x)^2 && (20)\\ &\quad +(z_x \sin\tilde{\phi} + z_y \cos\tilde{\phi} - \delta_y)^2 && (21)\end{aligned}$$

It is easy to prove that the function $f(\tilde{\phi}, \delta_x, \delta_y)$ is not convex. But it can be seen clearly that when $\tilde{\phi}$ (the error of robot orientation) is close to zero, the function is close to a quadratic thus convex function.

Moreover, when considering the sum of the terms, the non-convexity of the feature observations can probably be compensated by convexity in other terms of (9) in particular in the linear orientation part of odometry (10). More work is necessary to further investigate the near-convexity of the objective function (9).

V. A NEW STATE VECTOR FOR LEAST SQUARES SLAM

In this paper, we propose to use a state vector as "relative" as possible in SLAM. The new state vector is given by

$$\begin{aligned} X_{rel} = &(X_{r_1}^{r_0}, \phi_{r_1}^{r_0}, X_{r_2}^{r_1}, \phi_{r_2}^{r_1}, \cdots, X_{r_p}^{r_{p-1}}, \phi_{r_p}^{r_{p-1}}, \\ &X_{f_1}^{r_{m_1}}, \cdots, X_{f_N}^{r_{m_N}}). \end{aligned} \quad (22)$$

Here $r_{m_j} (1 \le j \le p)$ is the robot pose when the feature f_j is first observed. Note here that it is possible that $r_{m_i} = r_{m_j}$ for some $i \neq j$. An example of the state vector is illustrated in Figure 3(b).

A. Least squares formulation

Using the new state vector, the least squares problem becomes to minimize

$$\begin{aligned} &\sum_{j=1}^{p} (O_j^{j-1} - H_{rel}^{Oj}(X_{rel}))^T P_{Oj}^{-1} (O_j^{j-1} - H_{rel}^{Oj}(X_{rel})) \\ &+ \sum_{i,j} (Z_j^i - H_{rel}^{Z_j^i}(X_{rel}))^T P_{Z_j^i}^{-1} (Z_j^i - H_{rel}^{Z_j^i}(X_{rel})) \end{aligned} \quad (23)$$

where X_{rel} is defined in (22). For this formulation, all the functions $H_{rel}^{Oj}(X_{rel})$ and part of the functions $H_{rel}^{Z_j^i}(X_{rel})$ are simple linear functions. Using this new formulation, the quadratic part and non-quadratic part of the objective function are clearly distinguished.

B. Odometry information function $H_{rel}^{Oj}(X_{rel})$

Since relative pose $(X_{r_j}^{r_{j-1}}, \phi_{r_j}^{r_{j-1}})$ is part of the state vector, the odometry function is a very simple linear function

$$H_{rel}^{Oj}(X_{rel}) = (X_{r_j}^{r_{j-1}}, \phi_{r_j}^{r_{j-1}}). \quad (24)$$

C. Observation function $H_{rel}^{Z_j^i}(X_{rel})$ for first observations

When the feature is observed first time, the observation function is also a very simple linear function. In fact, for observation made from r_{m_j} to f_j, the observation function is

$$H_{rel}^{Z_j^{m_j}}(X_{rel}) = X_{f_j}^{r_{m_j}}. \quad (25)$$

D. Observation function $H_{rel}^{Z_j^i}(X_{rel})$ for subsequent observations

When the feature is observed second (or more) time, the observation function depends on when the same feature was observed first time. It is a function of a number of odometry and its first observation given by

$$H_{rel}^{Z_j^i}(X_{rel}) = X_{f_j}^{r_i} = (R(\phi_{r_i}^{r_{m_j}}))^T (X_{f_j}^{r_{m_j}} - X_{r_i}^{r_{m_j}}) \quad (26)$$

where

$$\begin{aligned} \phi_{r_{m_j+2}}^{r_{m_j}} &= \phi_{r_{m_j+1}}^{r_{m_j}} + \phi_{r_{m_j+2}}^{r_{m_j+1}} \\ &\vdots \\ \phi_{r_i}^{r_{m_j}} &= \phi_{r_{m_j+1}}^{r_{m_j}} + \phi_{r_{m_j+2}}^{r_{m_j+1}} + \cdots + \phi_{r_{i-1}}^{r_{i-2}} + \phi_{r_i}^{r_{i-1}} \end{aligned} \quad (27)$$

and

$$\begin{aligned} X_{r_i}^{r_{m_j}} = &X_{r_{m_j+1}}^{r_{m_j}} + R(\phi_{r_{m_j+1}}^{r_{m_j}}) X_{r_{m_j+2}}^{r_{m_j+1}} \\ &+ R(\phi_{r_{m_j+2}}^{r_{m_j}}) X_{r_{m_j+3}}^{r_{m_j+2}} \\ &+ \cdots \\ &+ R(\phi_{r_{i-1}}^{r_{m_j}}) X_{r_i}^{r_{i-1}}. \end{aligned} \quad (28)$$

Using the fact that $R(\theta_1 + \theta_2) = R(\theta_1)R(\theta_2)$, we have

$$\begin{aligned} &X_{r_i}^{r_{m_j}} \\ = &X_{r_{m_j+1}}^{r_{m_j}} + R(\phi_{r_{m_j+1}}^{r_{m_j}})[X_{r_{m_j+2}}^{r_{m_j+1}} + R(\phi_{r_{m_j+2}}^{r_{m_j+1}}) \\ &\cdot (X_{r_{m_j+3}}^{r_{m_j+2}} + \cdots + R(\phi_{r_{i-1}}^{r_{i-2}}) X_{r_i}^{r_{i-1}})]. \end{aligned} \quad (29)$$

E. Pros and cons of the new state vector

The only nonlinear part of the objective function in (23) is the subsequent observations to features as shown in (26). The two major advantages by using relative state vector as comparing to using absolute state vector are: (i) the odometry information function is completely linear; (ii) the nonlinearity of the observation function now depends on the accumulated robot orientation error from the robot pose when the feature is first observed to the current pose, instead of the accumulated robot orientation error form the first pose to the current pose. So the accumulated error and the potential nonlinearity/nonconvexity is reduced.

The side effect of this new state vector is that the information matrix is not as sparse as that of the original formulation, especially when there are a lot of loop closure in the robot trajectory. This may increase the computational cost of solving the least squares problem (23).

VI. WHAT IS THE ADVANTAGE OF USING MAP JOINING?

Recently we have shown in [11] that map joining problem can be formulated as an optimization problem where each local map is regarded as an integrated observation. We will show that by doing so, the nonlinearity involved is reduced significantly.

A. Reduced nonlinearity in local map building

When small local maps are built by least squares approach, the nonlinearity involved is less than that of building a large map. This is due to the smaller accumulated error of the robot poses as shown $\phi_{r_i}^{r_0}$ in (11) and $\phi_{r_i}^{r_{m_j}}$ in (26).

B. Map joining using absolute state vector

Suppose there are k local maps. The state vector of Iterative Sparse Local Submap Joining (I-SLSJF) [11] contains all the feature positions and robot end poses of each local map:

$$X_{join} = (X_{r_{1e}}^{r_0}, \phi_{r_{1e}}^{r_0}, \cdots, X_{r_{ke}}^{r_0}, \phi_{r_{ke}}^{r_0}, X_{f_1}^{r_0}, \cdots, X_{f_N}^{r_0}) \quad (30)$$

where r_{je} is the robot end pose of local map j ($1 \leq j \leq k$).

Suppose local map j is given by $(\hat{X}_j^L, P_j^L)$ as in a traditional EKF SLAM. Also suppose the features involved in local map j are $f_{j1}, \cdots, f_{jn}$, then the local map state estimate $\hat{X}_j^L$ can be regarded as an observation of the true relative positions from the robot start pose $X_{r_{(j-1)e}}^{r_0}, \phi_{r_{(j-1)e}}^{r_0}$ to the features $X_{f_{j1}}^{r_0}, \cdots, X_{f_{jn}}^{r_0}$ and the robot end pose $X_{r_{je}}^{r_0}, \phi_{r_{je}}^{r_0}$. That is,

$$\hat{X}_j^L = H_j(X_{join}) + w_j \quad (31)$$

where

$$H_j(X_{join}) = \begin{pmatrix} R(\phi_{r_{(j-1)e}}^{r_0})(X_{r_{je}}^{r_0} - X_{r_{(j-1)e}}^{r_0}) \\ \phi_{r_{je}}^{r_0} - \phi_{r_{(j-1)e}}^{r_0} \\ R(\phi_{r_{(j-1)e}}^{r_0})(X_{f_{j1}}^{r_0} - X_{r_{(j-1)e}}^{r_0}) \\ \vdots \\ R(\phi_{r_{(j-1)e}}^{r_0})(X_{f_{jn}}^{r_0} - X_{r_{(j-1)e}}^{r_0}) \end{pmatrix}$$

and w_j is the zero-mean Gaussian "observation noise" whose covariance matrix is P_j^L (when $j = 1$, $X_{r_{(j-1)e}}^{r_0} = [0,0]^T, \phi_{r_{(j-1)e}}^{r_0} = 0$).

So the problem of fusing local maps 1 to k is to estimate the global state X_{join} using all the local map information (31) for $j = 1, \cdots, k$. This problem can be formulated as a least squares problem. That is, finding X_{join} to minimize

$$\sum_{j=1}^{k} (\hat{X}_j^L - H_j(X_{join}))^T (P_j^L)^{-1} (\hat{X}_j^L - H_j(X_{join})). \quad (32)$$

C. Map joining using relative state vector

Relative state vector can also be used in the map joining step. The state vector is given by

$$\begin{aligned} X_{join,rel} \;=\; & (X_{r_{1e}}^{r_0}, \phi_{r_{1e}}^{r_0}, X_{r_{2e}}^{r_{1e}}, \phi_{r_{2e}}^{r_{1e}}, \cdots, X_{r_{ke}}^{r_{(k-1)e}}, \\ & \phi_{r_{ke}}^{r_{(k-1)e}}, X_{f_1}^{r_{m_1 e}}, \cdots, X_{f_N}^{r_{m_N e}}) \end{aligned} \quad (33)$$

where $r_{m_j e}$ is the robot end pose of local map m_j if f_j is first observed in local map $m_j + 1$.

Suppose $f_{j1}, \cdots, f_{jl}$ are first observed in some previous local maps while $f_{j(l+1)}, \cdots, f_{jn}$ are first observed in local map j, then the local map information can be expressed as

$$\hat{X}_j^L = H_{j,rel}(X_{join,rel}) + w_j \quad (34)$$

where

$$H_{j,rel}(X_{join,rel}) = \begin{pmatrix} X_{r_{je}}^{r_{(j-1)e}} \\ \phi_{r_{je}}^{r_{(j-1)e}} \\ R(\phi_{r_{(j-1)e}}^{r_{m_{j_1} e}})(X_{f_{j1}}^{r_{m_{j_1} e}} - X_{r_{(j-1)e}}^{r_{m_{j_1} e}}) \\ \vdots \\ R(\phi_{r_{(j-1)e}}^{r_{m_{j_l} e}})(X_{f_{jl}}^{r_{m_{j_l} e}} - X_{r_{(j-1)e}}^{r_{m_{j_l} e}}) \\ X_{f_{j(l+1)}}^{r_{(j-1)e}} \\ \vdots \\ X_{f_{jn}}^{r_{(j-1)e}} \end{pmatrix} \quad (35)$$

and w_j is the zero-mean Gaussian "observation noise" whose covariance matrix is P_j^L.

So the problem of fusing local maps 1 to k is to find the state $X_{join,rel}$ to minimize

$$\sum_{j=1}^{k} (\hat{X}_j^L - H_{j,rel}(X_{join,rel}))^T (P_j^L)^{-1} (\hat{X}_j^L - H_j(X_{join,rel})). \quad (36)$$

D. Advantages of map joining

Using map joining, not only the number of poses but also the degree of nonlinearity are significantly reduced. Especially when the relative state vector is used, it can be seen from (35) that the non-quadratic part is very small (only for f_{j_1} to f_{j_l}). This significantly reduce the nonconvexity of the problem. Moreover, the density of non-zero elements in the information matrix is small as compared with that of a single map SLAM using relative state vector, because the loop closing only take place at the local map level.

VII. Related work and Discussions

A few years back, the work by Olson et. al [8] surprised many SLAM researchers including us. How come stochastic gradient approach works so well for SLAM? The Tree-based network optimizer (TORO) algorithm by Grisetti et. al [7] further demonstrated very promising results where very large scale problems can be solved very efficiently without the need of good initial values. Especially when the covariance matrix of the relative pose is close to spherical [9]. This made us more curious about the special underlining structure of SLAM. SLAM must be a very special optimization problem!

Rizzini is probably another researcher who has noticed this fact and investigated into it [12]. In his work, he focused on the trajectory based SLAM and aimed at finding closed-form solutions of stationary points (local minima). In our work, we are trying to prove the near convexity of feature based SLAM problem, which we believe to be more general than the trajectory based SLAM problem.

The initial analysis on the convexity in this paper confirms that the accumulated orientation error is a key factor that governing the nonlinearity and nonconvexity of the SLAM problem. This is coincide with some recent research [13][14] where robot orientation error is shown to be the main cause of SLAM inconsistency. The benefit of having spherical covariances is due to the cancellation of highly nonlinear terms introduced by $R(\phi)$ as shown in (16). This explains why TORO can only performs well and why both the Victoria Park data and the DLR data set have the "magic" convergence property, all with spherical covariances.

The use of relative state vector in SLAM is not new. For example, it is shown in [9] that using a relative state vector makes the proposed constraint network optimization algorithm perform extremely well. Relative state vector has also been used in D-SLAM [16]. In [17], relative robot pose estimate is proposed to be used as a performance metric to compare different trajectory based SLAM algorithm. So why not directly using relative poses as state vector in the estimation/optimization?

In [10], SLAM was described as "Certainty of relations despite uncertainty in positions". If relative state vector (relation) is used, then the SLAM problem becomes more "certain". After a good and consistent estimate of the "relations" is obtained, to compute the "positions" is trivial.

Map joining has already been demonstrated to be an efficient strategy for large-scale SLAM [4]. It is also commented that map joining can reduce linearization error [15]. This paper further confirms this by showing how the nonlinearity and nonconvexity is reduced by using map joining.

VIII. Conclusions and Future Works

This paper provides some evidence of the underlining special structure of feature based SLAM problems. Some initial analysis on the convexity of SLAM problem is performed. The non-quadratic terms of SLAM optimization problem is further clearly distinguished by using a relative state vector. Moreover, how map joining can reduce the nonconvexity and nonlinearity is clearly explained.

The results in this paper clearly show that SLAM is a special optimization problem. The analysis of the convexity and nonlinearity helps to explain some unbelievable results in SLAM. Further more rigorous investigation of the underlining special structure in SLAM is necessary and will benefit SLAM community significantly. For example, more efficient and reliable SLAM algorithms could be developed by using the special structure of the problem. The work in these directions is underway.

References

[1] T. Bailey and H. Durrant-Whyte, "Simultanouse localization and mapping (SLAM): Part II". *IEEE Robotics & Automation Magazine*, vol. 13, Issue 3, Sept. 2006, pp. 108-117.

[2] M. Agrawal, K. Konolige, "FrameSLAM: From bundle adjustment to real-time visual mapping". *IEEE Transactions on Robotics*, Vol. 24, No. 5. 1066-1077, October 2008.

[3] F. Dellaert and M. Kaess, "Square root SAM: Simultaneous localization and mapping via square root information smoothing". *International Journal of Robotics Research*, vol. 25, no. 12, December 2006, pp. 1181-1203.

[4] S. Huang, Z. Wang and G. Dissanayake. "Sparse local submap joining filter for building large-scale maps". *IEEE Transactions on Robotics*, 2008, Vol. 24, No. 5, 1121-1130, October 2008.

[5] J. E. Guivant and E. M. Nebot, "Optimization of the simultaneous localization and map building (SLAM) algorithm for real time implementation," *IEEE Transactions on Robotics and Automation*, vol. 17, pp. 242-257, 2001.

[6] J. Kurlbaum, U. Frese, "A benchmark data set for data association". Technical Report, University of Bremen, available online: http://www.sfbtr8.uni-bremen.de/reports.htm Data available on http://radish.sourceforge.net/

[7] G. Grisetti, D. L. Rizzini, C. Stachniss, E. Olson and W. Burgard, "Online constraint network optimization for efficient maximum likelihood mapping". *IEEE International Conference on Robotics and Automation*, Pasadena, California, May 19-23, 2008, pp. 1880-1885.

[8] E. Olson, J. Leonard, and S. Teller. "Fast iterative optimization of pose graphs with poor initial estimates". *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2262-2269, 2006.

[9] G. Grisetti, C. Stachniss, and W. Burgard, "Non-linear constraint network optimization for efficient map learning", *Transations on Intelligent Transportation Systems.* Volume 10, Issue 3, Pages 428-439, 2009

[10] U. Frese. "A discussion of simultaneous localization and mapping". *Autonomous Robots*, Volume 20, Issue 1, (January 2006), pp. 25-42.

[11] S. Huang, Z. Wang, G. Dissanayake, and U. Frese, "Iterated SLSJF: A sparse local submap joining algorithm with improved consistency", *2008 Australiasan Conference on Robotics and Automation.* Canberra, Australia, December 2008. Available online: http://www.araa.asn.au/acra/acra2008/papers/pap102s1.pdf

[12] D. L. Rizzini, "Towards a closed-form solution of constraint networks for Maximum Likelihood mapping", *the 14th International Conference on Advanced Robotics (ICAR 2009)*, Munich, Germany, from June 22th to 26th, 2009.

[13] T. Bailey, J. Nieto, J. Guivant, M. Stevens and E. Nebot. "Consistency of the EKF-SLAM algorithm". *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3562-3568, Beijing, China, October 9-15, 2006.

[14] S. Huang and G. Dissanayake, "Convergence and consistency analysis for Extended Kalman Filter based SLAM". *IEEE Transactions on Robotics*, 2007, vol. 23, no. 5, 1036-1049.

[15] J. D. Tardos, J. Neira, P. M. Newman and J. J. Leonard, "Robust mapping and localization in indoor environments using sonar data", *International Journal of Robotics Research*, vol. 21, no. 4, April 2002, pp. 311-330.

[16] Z. Wang, S. Huang and G. Dissanayake, "Decoupling localization and mapping in SLAM using compact relative maps", *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Edmonton, Alberta, Canada, August 2-6, 2005. 1041-1046.

[17] R. Kummerle, B. Steder, C. Dornhege, M. Ruhnke, G. Grisetti, C. Stachniss, A. Kleiner, "On measuring the accuracy of SLAM algorithms", *Autonomous Robots*, (2009) 27: 387-407.