Photometric Based Loop Closure Correction and Pose Graph Optimization With Non-linear Factor Recovery

1st Xingwe Qu
Munich, Germany
xingwei@in.tum.de

2nd Nikolaus Demmel
Munich, Germany
demmeln@in.tum.de

3rd Prof. Dr.-Ing. Klaus Diepold
Munich, Germany
kldi@tum.de

4rd Prof. Dr. Daniel Cremers
Munich, Germany
cremers@.in.tum.de

Abstract—The purpose of this research work is to investigate some information matrix into PGO (pose graph optimization). Considering the fundamental challenges associated with this problem, such as the null-space nature of SLAM system and the ambiguity of the scale drift during monocular visual odometry. NFR(Nonlinear Factor Recovery) was introduced to extract an information matrix from a rank deficient marginalized hessian matrix such that a set of non-linear factors was reconstructed that make an optimal approximation of the information on the pose graph. Then, we proposed a photometric based $\text{Sim}(3)$ relative pose estimation. In experiments on a public benchmark, we demonstrate a competitive performance of our method over the LDSO original approach.

Index Terms—SLAM, Pose-graph Optimization

I. INTRODUCTION

The objective of pose-graph optimization is to estimate a set of camera poses (poses and orientatnions) from relative pose measurements. The problem can be formulated as a least squares minimization. However, due to the unknown variance of each poses, this problem usually initialize each state with the same weight. In a visual SLAM context, pose-graph optimization begins after the visual odometry, which could help us to extract information matrix of each pose.

In this paper, we extract non-linear factors from the marginalized keyframe at the font-end and use them as the initial guess of pose for global pose-graph optimization. The key idea, under the assumption of gaussian noise, the maximum likelihood estimation can be obtained by minizing the weigthed sum of residuals. the non-linear factor recovery can approximates a dense distribution stored in the original SLAM factor graph. Then we could transfer the information accumulated for our global pose graph optimization.

A. Contributions

(i) We demostrate that weighted global pose graph optimization can improve mapp accuracy significantly over pose-graph optimization. (ii) we propose a photometric based relative pose estimation and evaluate as well as accuracy. (iii) we improve the LDSO loop closure detection strategy which can improve the robustness of this system.

II. RELATED WORK

A. Graph-based SLAM

Graph-based SLAM maintains a global graph whose nodes represent camera's poses or landmarks and an edge represents a sensor measurement that constrains the connected pose.

Once such a graph is constructed, SLAM uses graph optimization method (i.e. nonlinear least-squares error minimization via the Gauss-Newton or Levenberg-Marquardt algorithm) to find a configuration of the nodes that is maximally consistent with all the constraints. The graph optimization procedure, with the presence of both camera pose and landmarks in the graph, is called Bundle-Adjustment (BA). [1] Monocular SLAM systems that apply graph optimization include PTAM [2], LSD-SLAM [3], DSO [4], etc

In monocular SLAM, loop closure is solved through a pose graph optimization with 7DoF similarity constraints (Sim(3)) to correct the scale drift. A pose graph is built on selected key-frame connected by the pose-pose constraints. Pose-pose constraints are defined by covisiblity and estimated by a geometry base VO front end. Two poses are connected to each other if they share enough common features. The graph built on covisiblity is called Covisiblity Graph. In order to achieve scalable, real-time performance, Ral et al. [5] proposed to perform loop closing on a much lighter Essential Graph which retains all the nodes (key-frames) from Covisibility Graph and a subset of edges with high covisibility. As shown in [5], the optimization of a properly constructed Essential Graph is already very accurate that full Bundle Adjustment only makes marginal improvement. We take the idea of loop closing with graph optimization and apply it to an unsupervised learning based visual odometry.

Loop closure is triggered by the place recognition technique. Appearance or image-image matching based methods, such as the bag of word approaches FAB-MAP [6] and DBoW2 [7], are dominating this area for their high efficiency. Ral Mur-Artal et al. [7] [5] proposed a bag of words place recognizer built on DBoW2 with ORB feature and successful achieved real-time loop closing. Then, Xiang et al. [8] use DBoW3 to give the DSO [4] a loop closure ability. In this work, for efficiency and simplicity, we use LDSO [8] as an initialization for our further optimization.

B. Non-linear Factor Recovery

Non-linear factor recovery is a general framework to marginalize the node in a graph-based SLAM system. edge pruning for a graph-based SLAM was first proposed in 2011 by kretzhschmar et al. [9], where depicts an example factor graph input, together with one of node removed. They further employed the Chow-Liu tree approximation [10],

III. PRELIMINARIES

In this paper, we write matrices as bold capital letters (e.g. $\mathbf{T}$) and vectors as bold lowercase letters (e.g. $\boldsymbol{\xi}$). Rigid-body poses are represented as $(\mathbf{R}, \mathbf{p}) \in SO(3) \times \mathbb{R}^3$ or as transformation matrices $\mathbf{T} \in SE(3)$ and $\mathbf{S} \in Sim(3)$ when needed. Incrementing a rotation $\mathbf{R}$ by an increment is defined as $\mathbf{R} \oplus \boldsymbol{\xi} = \text{Exp}(\boldsymbol{\xi})\mathbf{R}$. The difference between two rotations R and R is calculated as $\mathbf{R}_1 \ominus \mathbf{R}_2 = \text{Log}(\mathbf{R}_1 \mathbf{R}_1^{-1})$ such that $(\mathbf{R} \oplus \boldsymbol{\xi}) \ominus \mathbf{R} = \boldsymbol{\xi}$. Here we use Exp: $\mathbb{R}^3 \rightarrow SO(3)$, which is a composition of the hat operator ($\mathbb{R}^3 \rightarrow \mathfrak{so}(3)$)and the matrix exponential ($\mathfrak{so}(3) \rightarrow SO(3)$) and maps rotation vectors to their corresponding rotation matrices, and its inverse Log:$SO(3) \rightarrow \mathbb{R}^3$. For all other variables, such as translation , scale, we define $\oplus$ and *ominus* as regular addition and subtraction.

In the following we will use a state $\mathbf{s}$ that is defined as a tuple of several rotation, translation and scale variables, and a function $\mathbf{r}(\mathbf{s})$ that depends on it and can also produce rotations ,translation and scale as the result. An increment $\boldsymbol{\xi}$ Rn is a stacked vector with all the increments of the variables in $\mathbf{s}$. Then, the Jacobian of the function with respect to the increment is defined as

$$\mathbf{J}_{\mathbf{r}(\mathbf{s})} = \lim_{\boldsymbol{\xi} \rightarrow 0} \frac{\mathbf{r}(\mathbf{s} \oplus \boldsymbol{\xi} \ominus \mathbf{r}(\mathbf{s})}{\boldsymbol{\xi}} \quad (1)$$

Here, $\mathbf{s} \oplus \boldsymbol{\xi}$ denotes that each component in s is incremented with the corresponding segment in $\boldsymbol{\xi}$ using the appropriate definition of the $\oplus$ opertor, and simliarly for $\ominus$. The limit is done component-wise, such that the Jacobian is a matrix. For Euclidean quantities, this definition is just a normal derivative with an extension for rotations, both as function value and as function argument. For more details and possible alternative formulations we refer the reader to [11] [12]

We define the non-linear re-weighted least squares problems, we minimize functions of the form

$$E(\mathbf{s}) = \frac{1}{2}\mathbf{r}(s)^{\mathsf{T}}\mathbf{W}\mathbf{r}(\mathbf{s}), \quad (2)$$

which is a squared norm of the sum residuals with a information matrix $\mathbf{W}$. In this case, $\mathbf{r}((s))$ is purely vector-valued. Near the current state $\mathbf{s}$ we can use a linear approximation of the residual, which leads to

$$E(\mathbf{s} \oplus \boldsymbol{\xi}) = E(s) + \boldsymbol{\xi}^{\mathsf{T}}\mathbf{J}_{\mathbf{r}(\mathbf{s})}^{\mathsf{T}}\mathbf{W}\mathbf{r}(\mathbf{s}) + \frac{1}{2}\boldsymbol{\xi}^{\mathsf{T}}\mathbf{J}_{\mathbf{r}(\mathbf{s})}^{\mathsf{T}}\mathbf{W}\mathbf{J}_{\mathbf{r}(\mathbf{s})}\boldsymbol{\xi} \quad (3)$$

The optimum of this approximated energy can be attained using the Gauss-Newton increment

$$\boldsymbol{\xi}^{\star} = -(\mathbf{J}_{\mathbf{r}(\mathbf{s})}^{\mathsf{T}}\mathbf{W}\mathbf{J}_{\mathbf{r}(\mathbf{s})})^{-1}\mathbf{J}_{\mathbf{r}(\mathbf{s})}^{\mathsf{T}}\mathbf{r}(\mathbf{s}) \quad (4)$$

With this, we can iteratively update the state $\mathbf{s}_{i+1} = \mathbf{s} \oplus \boldsymbol{\xi}^{\star}$ until convergence.

IV. LOOP CLOSING CORRECTION WITH NON-LINEAR FACTOR RECOVERY

A. Loop Closing In DSO

Before delving into details of how our photometric based relative pose estimation works, we briefly present the general framework and formulation of LDSO [8]. LDSO is a extension of DSO [4] , where 8 keyframes are maintained and their states are jointly optimized in a sliding window. Keyframes $\mathcal{F}\{\mathbf{T}_1, ..., \mathbf{T}_m\}$ with SE(3) poses and points $\mathcal{P}\{\mathbf{p}_1, ..., \mathbf{p}_m\}$ inverse depth d are optimized through the photometric error which is defined as [4]:

$$C_{full} := \sum_{i \in \mathcal{F}} \sum_{j \in \mathcal{F}} \sum_{k \in \mathcal{P}} C_{i,j,k}(\mathbf{R}_i, \mathbf{t}_i, a_i, b_i, d_k, \mathbf{R}_j, \mathbf{t}_j, a_j, b_j),$$
$$(5)$$

$$where C_{i,j,k}(\cdot) := \sum_{p \in \mathcal{P}} \omega_{\mathbf{p}}\|(I_j(\mathbf{p}') - b_j) - \frac{t_j e^{a_j}}{t_i e^{a_i}}(I_i(\mathbf{p} - b_i))\|_{\gamma}$$
$$(6)$$

$$\omega_{\mathbf{p}} := \frac{c^2}{c^2 + \|i(\mathbf{p})\|^2}, \quad p' := \Pi(\mathbf{R}_{i,j}\Pi^{-1}(\mathbf{p}, d_k) + \mathbf{t}_{i,j}) \quad (7)$$

where, a and b are the affine light transform parameters, t is the exposure time, I denotes an image and $w_{\mathbf{p}}$ is a heuristic weighting factor with a constant scalar c and the gradient norm. $\mathbf{p}'$ is the reprojected pixel of $\mathbf{p}$ on I_j calculted by reprojection function, with $\Pi : \mathbb{R}^3 \rightarrow \Omega$ the projection and $\Pi^{-1} : \Omega \times \mathbb{R} \rightarrow \mathbb{R}^3$ the back-projection function, $\mathbf{R}$ and $\mathbf{t}$ the realitve rigid body between the two frames calculated from $\mathbf{T}_i$ and $\mathbf{T}_j$, d the inverse depth for each point.

Loop closure detection strategy follows from LDSO [8]. In an loop closure setup, LDSO first compute the descriptors for each keyframe and point, which are fed into the DBow3 database. Loop candidate is proposed for the current keyframe by querying the database. In LDSO, the best candidate is chosen for the loop closure relative pose calculation. Next, a Sim(3) transformation is optimized with Gauss-Newton method by minimizing the 3D and 2D geometric constraints. The cost function is:

$$\mathbf{E}_{loop} = \sum_{\mathbf{q}_i \in \mathcal{Q}_1} \left\|\mathbf{S}_{cr}\Pi^{-1}(\mathbf{p}_i, d_{\mathbf{p}_i}) - \Pi^{-1}(\mathbf{q}_i, d_{\mathbf{q}_i})\right\|_2 +$$
$$\sum_{\mathbf{q}_j \in \mathcal{Q}_2} \left\|\Pi(\mathbf{S}_{cr}\Pi^{-1}(\mathbf{p}_j, d_{\mathbf{p}_j}) - \mathbf{q}_j\right\|_2, \quad (8)$$

where $\mathcal{P} = \mathbf{p}_i$ is the reconstructed features in the loop candidate and $d_{\mathbf{p}_i}$ is their inverse depth. $\mathcal{Q} = \mathbf{q}_i$ is the matched features in the current keyframe, and $d_{\mathbf{q}_i}$ the sparse inverse depth of the current keyframes. Let $\mathcal{Q}_1 \subseteq \mathcal{Q}$ be those without depth and $\mathcal{Q}_2 = \mathcal{Q} - \mathcal{Q}_1$ be those with depth and Π and Π^{-1} are the projection and back projection functions.

In the end, a global pose graph optimization is applied to close long-terms loops for DSO. Although the pose graph optimization is light-weighted, and fast to compute the loop closure, the estimation still inevitably drift, and sometimes the system is unable to close loop.

It is worthing noting that even feature based loop closure correction, loop closure edges are still not sufficient for global pose graph optimization, which leads to inconsistent drift. To avoid these problem, we turn to the idea of using photometric based approach to estimate the relative pose and make the pose graph more accurate , which leaves us several other challenges: (i) How to use all the points in the selected keyframes to calculate the relative pose, instead of optimizing filtered matched feature points. One step further, some loop closure edges which could offer us reliable information are filtered during the feature based geometric error calculation, because of the insufficiency of the feature matching points. (ii) How to estimate the scale drift with only photometric based error. According to the LDSO's formulation , 3D constraints are involved to give the prior of scale drift. However, feature correspondences are needed for building the 3D constraints. How to present all the map points to scale estimation, which is challenging in this case as insufficient matched points for 3D constraints. (iii) Once the transformation between loop candidates are proposed, global pose graph optimization are applied for the whole system. How to set the information matrix between different loop closure edges.

Take these challenges into account we design our loop closing module as depicted in FigIV-A. In order to initialize our global Pose Graph Optimization and select a meaningful set of keyframes, landmarks and observations, we realize our implementation as a post-processing step of direct SLAM LDSO. During the loop-closure detection part, we soften the candidate keyframe selection strategy, and use photometric based relative pose estimation to get a Sim(3) transfomation between loop candidates. Finally, we refine global camera poses with non-linear factor recovery.

B. RELATIVE POSE CALCULATION WITH PHOTOMETRIC ERROR

In order to benefit from global pose graph optimization, we soften the loop closure selection strategy in LDSO. All the candidates which share enough similarity score in DBow3 become our relative pose estimation inputs. Then, we minimize the following cost function:

$$
\mathbf{E}_{loop} = \sum_{\mathbf{p}_j \in \mathcal{P}_1} \sum_{\mathbf{p}_i \in \mathcal{P}} \left\| \mathbf{I}(\Pi(\mathbf{S}_{cr}\Pi^{-1}(\mathbf{p}_i, d_{\mathbf{p}_j}) - \mathbf{I}(\mathbf{p}_i)) \right\|_2 +
$$
$$
\sum_{\mathbf{p}_j \in \mathcal{P}_2} \sum_{\mathbf{p}_j \in \mathcal{P}} \left\| \mathbf{I}(\Pi(\mathbf{S}_{cr}^{-1}\Pi^{-1}(\mathbf{p}_j, d_{\mathbf{p}_j})) - \mathbf{I}(\mathbf{p}_j)) \right\|_2,
$$
$$
\tag{9}
$$

where $\mathcal{P}$ is the pattern shown in Fig IV-B. Instead of using the brightness transfer parameters in DSO, we use local sum of squared difference to minimize the influence of the

illumination changes during the long-term loop. $\mathcal{P}_1$ and $\mathcal{P}_2$ are the reference frame, which we select from the old frames, and the current frame. We project both points in reference frame and current frame to each other two times to minimize the photometric error of each projected points.

C. Non-linear Factor Recovery

Non-linear factor recovery approximates a dense distribution stored in the linearized Markov blanket of the original factor graph with a different set of non-linear factors that yield a sparse factor graph topology. By linearization of the residual function of a non-linear least-squares problem. we obtain a multivariate Gaussian distribution $p(s) \sim N(\mu_o, \mathbf{H}_o^{-1})$ where μ is the state estimation. we want to construct another distribution $q(s) \sim N(\mu_a, \mathbf{H}_a^{-1})$ that well approximates the original distribution with a sparser factor graph topology. We follow the NFR and minimize the Kullback-Leibler divergence(KLD) between the recovered distribution and the original distribution in [13].

When we marginalize out a keyframe from the original factor graph, we save the current linearization and marginalize out everything except the keyframe poses. This gives us a factor that densely connects all keyframe poses in the optimization window. We use it to recover non-linear factors between the marginalized keyframe and all other keyframes. We define the relative pose residual function same as [14]:

$$
\mathbf{r}_{rel}(\mathbf{r}, \mathbf{z}_{rel}) = \mathrm{Log}(\mathbf{z}_{rel}\mathbf{T}_i\mathbf{T}_j^{-1}) \tag{10}
$$

where $\mathbf{z}_{rel}$ is our virtual measurement from the estimated state at the time of linearization. We recover pairwise relative-pose factors between the keyframe. This gives us a full-rank invertable Jacobian $\mathbf{J}_r$,then we can use the KLD to recover the information matrices of the factors with that the following closed-form solution exists.

$$
\mathbf{H}_i = (\{\mathbf{J}_r\boldsymbol{\Sigma}_o\mathbf{J}_r^T\}_i)^{-1} \tag{11}
$$

where $_i$ denotes the corresponding diagonal block, and $\boldsymbol{\Sigma}_o = \mathbf{H}_o^{-1}$.

D. Handling Rank Deficient Information

In the section IV-C, we assumed the information matrix $\mathbf{H}_o$ to be invertible. Unfortunately, when dealing with problems sunch as variable marginalization. this is often not be the case [15]. From a SLAM perspective, this implies that the Markov blanket lacks a rigid transforamtion edge to the world reference frame and also the scale.

In such a senario, $\mathbf{H}_o$ has k null eigenvalues, where k is the nullspace of the SLAM system , e.g. 7 for LDSO. if $\mathbf{H}_o$ has n rows, the distribution of $p(s)$ is actually an $(n-k)$-dimensional multivariate normal embedded in a n-dimensional space. Therefore, we propose to project $p(s)$ and $q(s)$ onto the $(n-k)$-dimensional informative subspace to compare the resulting $(n-k)$-dimensional distributions.

We compute an $(n-k) \times n$ projection matrix $\mathbf{P}$ by stacking
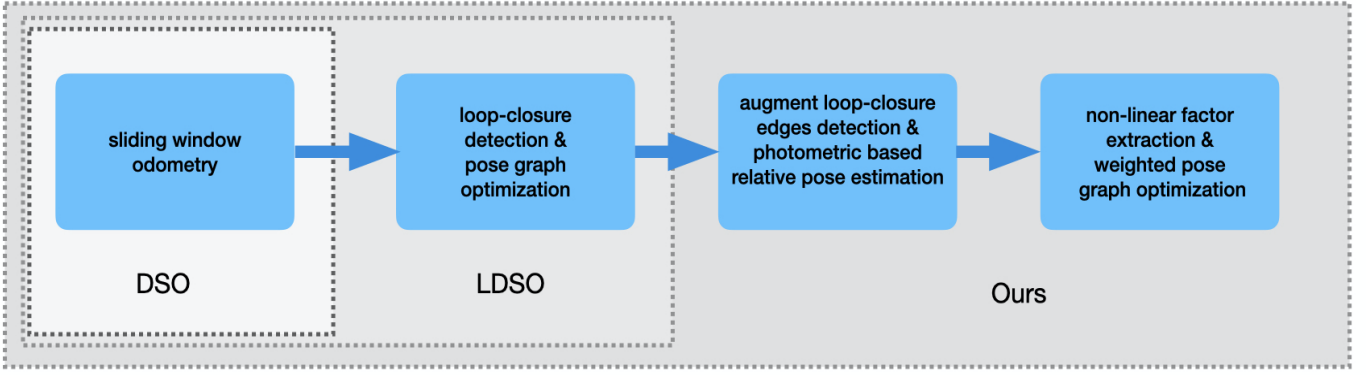
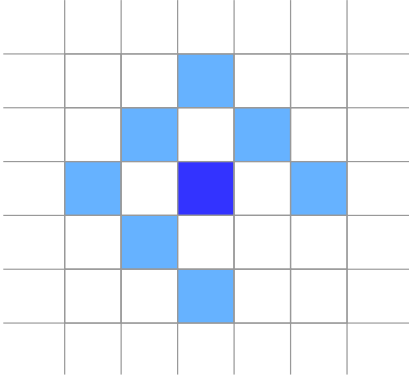Fig. 1: Pipline to construct and initializae the global PGO problem.

Fig. 2: Pattern for the local match.

the transpose of the eigen vectors of $\mathbf{H}_o$ corresponding to the nonzero eigenvalues. Since $\mathbf{H}_o$ is a symmetric real-valued matrix its eigen decomposition $\mathbf{H}_o = \mathbf{U}\Lambda\mathbf{U}^T$ is a real-valued and always exists:

$$\mathbf{H}_o = \mathbf{U}\Lambda\mathbf{U}^T = [\mathbf{U}_0 \ \ \mathbf{P}^T]\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Lambda_+ \end{bmatrix}\begin{bmatrix} \mathbf{U}_0^T \\ \mathbf{P} \end{bmatrix} \quad (12)$$

the projection matrix $\mathbf{P}$ acts as a linear operator to project any arbitrary information matrix $\mathbf{H}_a$ onto the lower-dimensional space by computing $\mathbf{P}\mathbf{H}_a\mathbf{P}^T$. To account for singular information matrices, we apply the following substitution into Kullback-Leibler divergence formula.

$$\mathbf{A} \rightarrow \mathbf{A}\mathbf{P}^T \quad (13)$$

$$\Sigma \rightarrow \Lambda_+^{-1} \quad (14)$$

This, however, does not guarantee that $q(x)$ is an $(n-k)$-dimensional distribution embedded in an n-dimensional space. We enforce this property by arifically limiting the rank of the Jacobian matrix $\mathbf{J}_r$. THe same efficient implementation for computation for computing the gradient can also be applied in the low rank case.

$$\mathbf{Y} = \mathbf{P}[\Lambda - (\mathbf{P}\mathbf{J}_r^T\mathbf{H}_o\mathbf{J}_r\mathbf{P}^T)^{-1}]\mathbf{P} \quad (15)$$

where $\mathbf{Y}$ represents the gradient. If the $\mathbf{J}_r$ is full rank column, the $\mathbf{H}_a$ can be derived by:

$$\mathbf{H}_i = (\{\mathbf{J}_r\mathbf{P}^T\Sigma_o\mathbf{P}^{-1}\mathbf{J}_r^T\}_i)^{-1} \quad (16)$$

V. EVALUATION

A. Setup

In order to evaluate the performance of the our weighted pose graph optimization. we evaluate our implementation on all eleven sequences from the EUROC MAV datasets. These constitute image sequences froma flying drone in an indoor setting. Accurate ground truth poses are available from motion capture or fusion of laser tracking and IMU. It worth noting that, since we use LDSO as initialization and baseline as our algorithm, we choose the same subset of frames for each sequences as in their evaluation and run our monocular pipline on images from the left camera.

B. Results

Accuracy In order to evaluate the accuracy of the keyframe poses after optimization, we employ the commly used Absolute Trajactory Error (ATE) in meters [16], which measures the root-mean-squared camera position error, after Sim(3) alignment of the estimated keyframe trajectory to the ground truth:

$$ATE = \min_{(\mathbf{R},(t),s)\in SO(3)\times\mathbb{R}^3\times\mathbb{R}} \sum_i \|\hat{\mathbf{x}}_i + s\mathbf{R}(\mathbf{R}_i^\mathsf{T})\mathbf{t}_i - \mathbf{t}\|_2 \quad (17)$$

where $\hat{\mathbf{x}}_i$ are ground truth camera postition and $-\mathbf{R}_i^\mathsf{T}\mathbf{t}_i$ are the correspoing estimates from global variables. Ground truth poses are linearly interpolated at image time stamps. Table 1 shows ATE for all sequences. Note that for MH04 and V203 the odoemtry has large scale/rotation drift and no verified loop closures where detected. In this failure case of the initialization global PGO does not help, however our weighted PGO can improves a little. Despite this, our weighted pgo can still improve the results comparing to the original PGO.
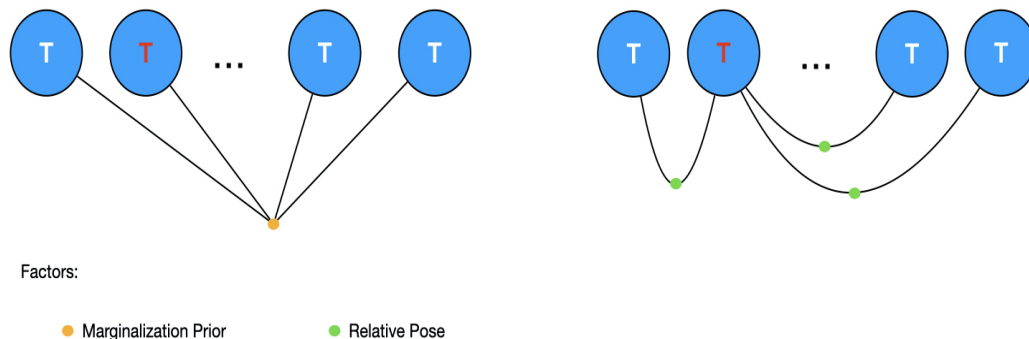
Fig. 3: Visualization of non-linear factor recovery. Left: Densely connected factor from marginalization saved from the LDSO before removing a keyframe pose. Right: Extracted non-linear factors that approximate the distribution stored in the original factor.

Sequence	MH01	MH02	MH03	MH04	MH05	V101	V102	V103	V201	V202	V203
odom	0.059	0.037	0.237	3.111	0.101	0.116	0.256	0.134	0.052	0.088	1.318
pgo	0.034	0.019	0.087	3.122	0.074	0.040	0.058	0.134	0.026	0.053	1.318
wpgo	0.02359	**0.02297**	0.02406	0.02347	0.02448	0.03045	0.05318				

TABLE I: ATE error for all Euroc sequences. The final ATE in meters for odometry (odom), pose-graph optimizaiton (pgo), our weighted pose-graph optimization shows that our approach significatly imporves over odometry and is competitive to the performance of pose-graph optimization

VI. CONCLUSION

In this report, we present a novel approach for $\text{Sim}(3)$ PGO that combines relative pose constraints in the sliding window and the information distribution of poses. We achieve this that successively recovers non-linear factors from LDSO estimate that summarize the relative pose constraints between keyframes. Considering the null-space of the monocular SLAM system, we use NFR to handle a rank deficient marginalized information matrix and test in Euroc dataset. The result is competitive comparing with the LDSO approach.

REFERENCES

[1] Y. Li, Y. Ushiku, and T. Harada, "Pose graph optimization for unsupervised monocular visual odometry," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 5439–5445, IEEE, 2019.

[2] G. Klein and D. Murray, "Parallel tracking and mapping on a camera phone," in *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pp. 83–86.

[3] J. Engel, T. Schöps, and D. Cremers, "Lsd-slam: Large-scale direct monocular slam," in *European conference on computer vision*, pp. 834–849, Springer, 2014.

[4] J. Engel, V. Koltun, and D. Cremers, "Direct sparse odometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Mar. 2018.

[5] M. J. M. M. Mur-Artal, Raúl and J. D. Tardós, "ORB-SLAM: a versatile and accurate monocular SLAM system," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[6] M. Cummins and P. Newman, "Fab-map: Probabilistic localization and mapping in the space of appearance," *The International Journal of Robotics Research*, vol. 27, no. 6, pp. 647–665, 2008.

[7] D. Gálvez-López and J. D. Tardós, "Bags of binary words for fast place recognition in image sequences," *IEEE Transactions on Robotics*, vol. 28, pp. 1188–1197, October 2012.

[8] X. Gao, R. Wang, N. Demmel, and D. Cremers, "Ldso: Direct sparse odometry with loop closure," in *iros*, October 2018.

[9] H. Kretzschmar, C. Stachniss, and G. Grisetti, "Efficient information-theoretic graph pruning for graph-based slam with laser range finders," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 865–871, IEEE, 2011.

[10] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE transactions on Information Theory*, vol. 14, no. 3, pp. 462–467, 1968.

[11] T. D. Barfoot, "State estimation for robotics,"

[12] M. Bloesch, H. Sommer, T. Laidlow, M. Burri, G. Nuetzi, P. Fankhauser, D. Bellicoso, C. Gehring, S. Leutenegger, M. Hutter, *et al.*, "A primer on the differential calculus of 3d orientations," *arXiv preprint arXiv:1606.05285*, 2016.

[13] M. Mazuran, W. Burgard, and G. D. Tipaldi, "Nonlinear factor recovery for long-term slam," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 50–72, 2016.

[14] V. Usenko, N. Demmel, D. Schubert, J. Stückler, and D. Cremers, "Visual-inertial mapping with non-linear factor recovery," *arXiv preprint arXiv:1904.06504*, 2019.

[15] M. Mazuran, G. D. Tipaldi, L. Spinello, and W. Burgard, "Nonlinear graph sparsification for slam,"

[16] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of rgb-d slam systems," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 573–580, IEEE, 2012.