A dissertation submitted to

TongjiUniversity in conformity with the requirements for

the degree of Masterof Engineering

Research and Implementation of Path Planning and Trajectory Tracking Algorithm for Monocular Vision Wheeled Robot

Candidate: Wang Ziqiang

Student Number: 1531651

School/Department: School of Electronics and InformationEngineering

Discipline: Engineering

Major: Control Engineering

Supervisor: AssociateProf. Xu Hegen

March, 2018

ABSTRACT

Warehouse logistics robots will work in different warehouse environments. In order to enable robots to perceive environment and plan path faster without modifying existing warehouses, we uses monocular camera to achieve an efficient robot system integration. Mapping and path planning the two main tasks presented in this paper. The direct method visual odometry is applied to localize, and the 3D position of major obstacles in the environment is calculated. We describe the terrain with occupied grid map , the 3D points are projected onto the robot motion plane, thus accessibility of each grid is determined. Based on the terrain information, the optimized A* algorithm is used for path planning. Finally, according to localization and planning, we control the robot track path. We also develop a path-tracking robot prototype. Simulation and experimental results verify the effectiveness and reliability of the proposed method.

Key Words: VSLAM, OGM, A*, Monocular vision, Mobile robot

Content

Chapter1 PHOTO RESIDUALS

1.1 INTRODUCTION

Windowed Optimization is a classic method in non-linear optimization.

1.1.1 NOTATION

Throughout the paper, we will write matrices as bold capital letters ($\mathbf{R}$) and vectors as bold lower case letters ($\boldsymbol{\xi}$), light lower-case letters to denote scalars (s). Light upper-case letters are used to represent functions (I).

Homogeneous camera calibration matrices are denoted by $\mathbf{K}$ as (2.1). Camera poses are represented by matrices of the special Euclidean group $\mathbf{T} \in SE(3)$, which transform a 3D coordinate from the camera coordinate system to the world coordinate system. In this paper, a homogeneous 2D image coordinate point $\mathbf{p}$ is represented by its image coordinate and inverse depth as (2.1) relative to its host keyframe I_i^L. The host keyframe is the frame the point got selected from. Corresponding homogeneous 3D world coordinate point $\mathbf{p}_w$ is denoted as (2.1). $\Pi_{\mathbf{K}}$ are used to denote camera projection functions. The jacobian of I_i, $\Pi_{\mathbf{K}}$ is denoted as (2.1)

$$
\begin{aligned}
&\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{K}^{-1} = \begin{pmatrix} f_x^{-1} & 0 & -f_x^{-1}c_x & 0 \\ 0 & f_y^{-1} & -f_y^{-1}c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\
&\mathbf{p} = \begin{pmatrix} u^i \\ v^i \\ 1 \\ d_{\mathbf{p}} \end{pmatrix}, \mathbf{p}_w = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, d_{\mathbf{p}} = z^{-1}, \mathbf{p} = d_{\mathbf{p}}\mathbf{K}\mathbf{p}_w = \Pi_{\mathbf{K}}(\mathbf{p}_w) \\
&\frac{\partial I_i^L(\mathbf{p})}{\partial \mathbf{p}} = (g_x, g_y, 0, 0), \frac{\partial \mathbf{p}}{\partial \mathbf{p}_w} = \begin{pmatrix} f_x z^{-1} & 0 & -x f_x z^{-2} & 0 \\ 0 & f_y z^{-1} & -y f_y z^{-2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & z^{-2} & 0 \end{pmatrix}
\end{aligned}
\tag{2.1}
$$

1.1.2 QUESTION IMPORT

Assume we observe 5 points $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5\}$ in 4 keyframes $\boldsymbol{\mathcal{F}} = \{I_1, I_2, I_3, I_4\}$, every keyframe has stereo vision (I_i^L, I_i^R) abbreviated as (iL, iR). A point can also be observed by other frame as shown in Table(2.1). Question is how to use Windowed Optimization method to make our observation more accurate ?

Table (2.1)

Image point	Host keyframe	Observe by
$\mathbf{p}_1$	$1L$	$1R, 2L$
$\mathbf{p}_2$	$2L$	$2R, 1L, 3L$
$\mathbf{p}_3$	$2L$	$3L$
$\mathbf{p}_4$	$3L$	$1L, 2L$
$\mathbf{p}_5$	$4L$	$3L, 4L$

1.2 SOLUTION

We use direct method to construct residual, Windowed Gauss-Newton method to optimization residual。

1.2.1 CONSTRUCT RESIDUAL

Dynamic multi-view stereo residuals $E_{ij}^{\mathbf{p}}$ are defined as

$$E_{ij}^{\mathbf{p}} = w_{\mathbf{p}} ||(r_{\mathbf{p}}^d)_{ij}||_\gamma, \quad (r_{\mathbf{p}}^d)_{ij} := I_j^L(\mathbf{p}^{'}) - b_j^L - \frac{e^{a_j^L}}{e^{a_i^L}}(I_i^L(\mathbf{p}) - b_i^L) \tag{2.2}$$

γ is Huber norm. a_i^L, b_i^L is affine brightness parameters to frame iL . $w_{\mathbf{p}}$ is a gradient-dependent weighting parameters, $\mathbf{p}$ in frame I_i^L projected to I_j^L is $\mathbf{p}^{'}$ as:

$$w_{\mathbf{p}} := \frac{c^2}{c^2 + ||\nabla I_i(\mathbf{p})||_2^2}, \quad \mathbf{p}^{'} = d_{\mathbf{p}}^{jL} \mathbf{K}(\mathbf{T}_j \mathbf{T}_i^{-1}((d_{\mathbf{p}}^{iL})^{-1} \mathbf{K}^{-1} \mathbf{p})) \tag{2.3}$$

Static one-view stereo residuals $E_{is}^{\mathbf{p}}$ are modified to

$$E_{is}^{\mathbf{p}} = w_{\mathbf{p}} ||r_{\mathbf{p}}^s||_\gamma, \quad r_{\mathbf{p}}^s := I_i^R(\mathbf{p}^{'}) - b_i^R - \frac{e^{a_i^R}}{e^{a_i^L}}(I_i^L(\mathbf{p}) - b_i^L) \tag{2.2}$$

Hostframe of $\mathbf{p}$ is I_i^L. a_i^R, b_i^R is affine brightness parameters to frame iR. $\mathbf{p}$ in frame I_i^L projected to I_i^R is $\mathbf{p}^{'}$ as :

$$\mathbf{p}^{'} = d_{\mathbf{p}}^{iR} \mathbf{K}(\mathbf{T}_{RL}((d_{\mathbf{p}}^{iL})^{-1} \mathbf{K}^{-1} \mathbf{p})) \tag{2.2}$$

Total residuals

$$E(\delta) = \sum_{i \in \mathcal{F}} \sum_{\mathbf{p} \in \mathcal{P}_i} \Big(\sum_{j \in obs^t(\mathbf{P})} E_{ij}^{\mathbf{p}} + \lambda E_{is}^{\mathbf{p}} \Big) \tag{2.2}$$

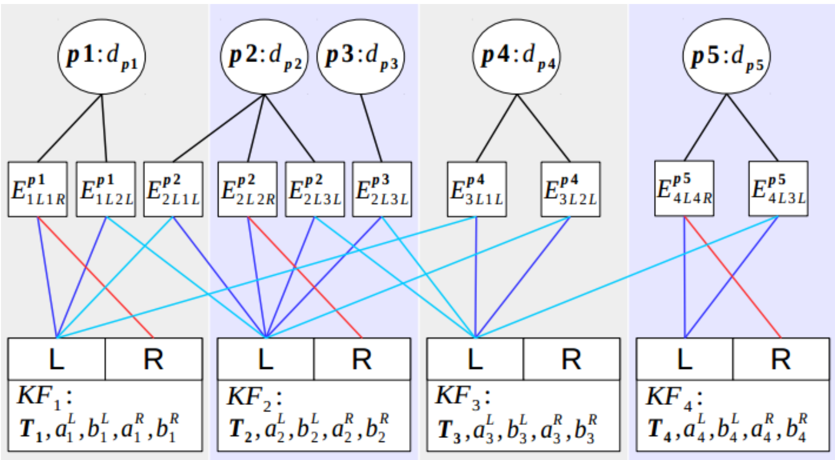
To balance the relative weights of temporal multi-view and static stereo, we introduce a coupling factor λ to weight the constraints from static stereo differently.

$$\boldsymbol{\delta}=\begin{pmatrix}(\boldsymbol{\xi}_1^T,\ldots,\boldsymbol{\xi}_{N_f}^T)^T\\(d_{\mathbf{p}_1},\ldots,d_{\mathbf{p}_{N_p}})^T\\(a_1^L,a_1^R,b_1^L,b_1^R)^T\\\vdots\\(a_{N_f}^L,a_{N_f}^R,b_{N_f}^L,b_{N_f}^R)^T\\(f_x,f_y,c_x,c_y)^T\end{pmatrix}\in\mathbb{R}^{10N_f+N_p+4},\boldsymbol{\xi}_i=(ln\mathbf{T}_i)^{\vee}\in\mathbb{R}^6 \tag{2.1}$$

$\mathcal{P}_i$ is a set of all image point host by frame iL. $obs^t(\mathbf{p})$ are the observations of $\mathbf{p}$ from temporal multi-view stereo. If there are N_p image point and N_f keyframes in $\mathcal{F}$, optimization variable $\boldsymbol{\delta}$ is

In this example, there are 7 dynamic residuals and 3 static residuals, Factor graph of the residuals function is

Total residuals is



$$\begin{aligned}E(\delta)&=E^{\mathbf{p}_1}_{1L2L}+E^{\mathbf{p}_2}_{2L1L}+E^{\mathbf{p}_2}_{2L3L}+E^{\mathbf{p}_3}_{2L3L}+E^{\mathbf{p}_4}_{3L1L}+E^{\mathbf{p}_4}_{3L2L}+E^{\mathbf{p}_5}_{4L3L}\\&\quad+E^{\mathbf{p}_1}_{1L1R}+E^{\mathbf{p}_2}_{2L2R}+E^{\mathbf{p}_5}_{4L4R}\\&=E_d(\delta)+E_s(\delta)\end{aligned}$$

$$E_s(\delta)=\begin{pmatrix}r^s_{\mathbf{p}_1}\\r^s_{\mathbf{p}_2}\\r^s_{\mathbf{p}_5}\end{pmatrix}^T\begin{pmatrix}\lambda w_{\mathbf{p}_1}&0&0\\0&\lambda w_{\mathbf{p}_2}&0\\0&0&\lambda w_{\mathbf{p}_5}\end{pmatrix}\begin{pmatrix}r^s_{\mathbf{p}_1}\\r^s_{\mathbf{p}_2}\\r^s_{\mathbf{p}_5}\end{pmatrix}=(\mathbf{r}^s)^T\mathbf{W}^s\mathbf{r}^s$$

$$\mathbf{J}_s=\begin{pmatrix}\frac{\partial r^s_{\mathbf{p}_1}}{\partial\boldsymbol{\xi}_1}\cdots&\frac{\partial r^s_{\mathbf{p}_1}}{\partial\boldsymbol{\xi}_4}&\frac{\partial r^s_{\mathbf{p}_1}}{\partial d_{\mathbf{p}_1}}\cdots&\frac{\partial r^s_{\mathbf{p}_1}}{\partial d_{\mathbf{p}_5}}&\frac{\partial r^s_{\mathbf{p}_1}}{\partial a_1^L}\cdots&\frac{\partial r^s_{\mathbf{p}_1}}{\partial b_4^R}&\frac{\partial r^s_{\mathbf{p}_1}}{\partial f_x}\cdots&\frac{\partial r^s_{\mathbf{p}_1}}{\partial c_y}\\\frac{\partial r^s_{\mathbf{p}_2}}{\partial\boldsymbol{\xi}_1}\cdots&\frac{\partial r^s_{\mathbf{p}_2}}{\partial\boldsymbol{\xi}_4}&\frac{\partial r^s_{\mathbf{p}_2}}{\partial d_{\mathbf{p}_1}}\cdots&\frac{\partial r^s_{\mathbf{p}_2}}{\partial d_{\mathbf{p}_5}}&\frac{\partial r^s_{\mathbf{p}_2}}{\partial a_1^L}\cdots&\frac{\partial r^s_{\mathbf{p}_2}}{\partial b_4^R}&\frac{\partial r^s_{\mathbf{p}_2}}{\partial f_x}\cdots&\frac{\partial r^s_{\mathbf{p}_2}}{\partial c_y}\\\frac{\partial r^s_{\mathbf{p}_5}}{\partial\boldsymbol{\xi}_1}\cdots&\frac{\partial r^s_{\mathbf{p}_5}}{\partial\boldsymbol{\xi}_4}&\frac{\partial r^s_{\mathbf{p}_5}}{\partial d_{\mathbf{p}_1}}\cdots&\frac{\partial r^s_{\mathbf{p}_5}}{\partial d_{\mathbf{p}_5}}&\frac{\partial r^s_{\mathbf{p}_5}}{\partial a_1^L}\cdots&\frac{\partial r^s_{\mathbf{p}_5}}{\partial b_4^R}&\frac{\partial r^s_{\mathbf{p}_5}}{\partial f_x}\cdots&\frac{\partial r^s_{\mathbf{p}_5}}{\partial c_y}\end{pmatrix}_{3\times 49} \tag{2.2}$$

$$E_d(\delta)=\begin{pmatrix}(r^d_{\mathbf{p}_1})_{12}\\(r^d_{\mathbf{p}_1})_{21}\\\vdots\\(r^d_{\mathbf{p}_5})_{43}\end{pmatrix}^T\begin{pmatrix}w_{\mathbf{p}_1}&0&\ldots&0\\0&w_{\mathbf{p}_1}&\ldots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\ldots&w_{\mathbf{p}_5}\end{pmatrix}\begin{pmatrix}(r^d_{\mathbf{p}_1})_{12}\\(r^d_{\mathbf{p}_1})_{21}\\\vdots\\(r^d_{\mathbf{p}_5})_{43}\end{pmatrix}=(\mathbf{r}^d)^T\mathbf{W}^d\mathbf{r}^d$$

$$\mathbf{J}_d = \begin{pmatrix} \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial \boldsymbol{\xi}_1} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial \boldsymbol{\xi}_4} & \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial d_{\mathbf{p}_1}} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial d_{\mathbf{p}_5}} & \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial a^L_1} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial b^R_4} & \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial f_x} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial c_y} \\ \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial \boldsymbol{\xi}_1} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial \boldsymbol{\xi}_4} & \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial d_{\mathbf{p}_1}} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial d_{\mathbf{p}_5}} & \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial a^L_1} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial b^R_4} & \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial f_x} \cdots & \frac{\partial (r^d_{\mathbf{p}_1})_{21}}{\partial c_y} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial \boldsymbol{\xi}_1} \cdots & \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial \boldsymbol{\xi}_4} & \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial d_{\mathbf{p}_1}} \cdots & \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial d_{\mathbf{p}_5}} & \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial a^L_1} \cdots & \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial b^R_4} & \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial f_x} \cdots & \frac{\partial (r^d_{\mathbf{p}_5})_{43}}{\partial c_y} \end{pmatrix}_{7\times 49} \tag{2.2}$$

Iteration δ^* can be calculated by

$$\begin{aligned} &(\mathbf{J}_s^T \lambda \mathbf{W}^s \mathbf{J}_s + \mathbf{J}_d^T \mathbf{W}^d \mathbf{J}_d)\delta^* = -(\mathbf{J}_s^T \lambda \mathbf{W}^s \mathbf{r}^s + \mathbf{J}_d^T \mathbf{W}^d \mathbf{r}^d) \\ &\mathbf{J}_s \in \mathbb{R}^{3\times 49}, \mathbf{W}^s \in \mathbb{R}^{3\times 3}, \mathbf{J}_s \in \mathbb{R}^{7\times 49}, \mathbf{W}^s \in \mathbb{R}^{7\times 7}, \end{aligned} \tag{2.2}$$

We construct residuals and its formulation.

1.2.2 JACOBIAN CITATION

We know for a Lie algebra $\boldsymbol{\rho} \in \mathbb{R}^3, \boldsymbol{\phi} \in \mathbb{R}^3, \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{pmatrix} \in \mathbb{R}^6$ and $\mathbf{p}_w$:

$$\begin{aligned} &\boldsymbol{\xi}^\wedge = \begin{pmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{pmatrix}^\wedge = \begin{pmatrix} \boldsymbol{\phi}^\wedge & \boldsymbol{\rho} \\ \mathbf{0}^T & \mathbf{0}^T \end{pmatrix} \in \mathbb{R}^{4\times 4} \\ &\boldsymbol{\epsilon} \in \mathbb{R}^3, \begin{pmatrix} \boldsymbol{\epsilon} \\ 1 \end{pmatrix}^\odot = \begin{pmatrix} \mathbf{E} & -\boldsymbol{\epsilon}^\wedge \\ \mathbf{0}^T & \mathbf{0}^T \end{pmatrix} \in \mathbb{R}^{4\times 6} \\ &\frac{\partial (exp(\xi^\wedge)\mathbf{p}_w)}{\partial \boldsymbol{\xi}} = \frac{\partial (\mathbf{T}\mathbf{p}_w)}{\partial \boldsymbol{\xi}} = (\mathbf{T}\mathbf{p}_w)^\odot \\ &\mathbf{T}\mathbf{p}_w = exp(\xi^\wedge)\mathbf{p}_w \approx (\mathbf{E} + \xi^\wedge)\mathbf{p}_w \\ &\frac{\partial (exp(\xi^\wedge)\mathbf{p}_w)}{\partial \boldsymbol{\xi}} \approx \frac{\partial (\mathbf{E} + \xi^\wedge)}{\partial \boldsymbol{\xi}} = \mathbf{0} + \frac{\partial (\xi^\wedge \mathbf{p}_w)}{\partial \boldsymbol{\xi}} \approx (\mathbf{T}\mathbf{p}_w)^\odot \\ &since, \frac{\partial (\mathbf{T}\mathbf{p}_w)}{\partial \boldsymbol{\xi}} = (\mathbf{T}^{-1}\mathbf{p}_w)^\odot = \frac{\partial (exp(-\xi^\wedge)\mathbf{p}_w)}{\partial \boldsymbol{\xi}} \\ &= \frac{\partial (\mathbf{E} - \xi^\wedge)}{\partial \boldsymbol{\xi}} = -(\mathbf{T}\mathbf{p}_w)^\odot \end{aligned} \tag{2.2}$$

1.2.3 JACOBIAN DERIVATION

1.2.3.1 Dynamic Parameter

Firstly, if $\mathbf{p}$ is neither observed by frame mL, mR nor hosted by nL, nR:

$$\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial \boldsymbol{\xi_m}} = \frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial \boldsymbol{\xi_n}} = \mathbf{0}^T, so \quad \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial \boldsymbol{\xi}_3} = \frac{\partial r^d_{(\mathbf{p}_1)}{}_{12}}{\partial \boldsymbol{\xi}_4} = \ldots = \mathbf{0}^T, \tag{2.2}$$

otherwise, we follow

$$\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial \boldsymbol{\xi_i}} = \frac{\partial (I^L_j(\mathbf{p}^{'}))}{\partial \boldsymbol{\xi_i}} = \frac{\partial (I^L_j(\mathbf{p}^{'}))}{\partial \mathbf{p}^{'}} \frac{\partial \mathbf{p}^{'}}{\partial \mathbf{p}^{'}_w} \frac{\partial \mathbf{p}^{'}_w}{\partial \boldsymbol{\xi_i}} \qquad (2.2)$$

$$\mathbf{p}^{'}_w = \mathbf{T}_j \mathbf{T}_i^{-1} \mathbf{p}_w = \mathbf{T}_j \mathbf{T}_i^{-1} ((d^{iL}_{\mathbf{p}})^{-1} \mathbf{K}^{-1} \mathbf{p})$$

For one frame iL, we have $\mathbf{p}$ and $\mathbf{K}$, then we can get

$$\left\{ \mathbf{p}_w = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} f_x^{-1} (d^{iL}_{\mathbf{p}})^{-1} (u^i - c_x) \\ f_y^{-1} (d^{iL}_{\mathbf{p}})^{-1} (v^i - c_y) \\ (d^{iL}_{\mathbf{p}})^{-1} \\ 1 \end{pmatrix} \right. \qquad (2.2)$$

$$\begin{aligned} \frac{\partial \mathbf{p}^{'}_w}{\partial \boldsymbol{\xi_i}} &= d^{jL}_{\mathbf{p}} \mathbf{K} \mathbf{T}_j \frac{\partial (\mathbf{T}_i^{-1} \mathbf{p}^{'}_w)}{\partial \boldsymbol{\xi_i}} = -d^{jL}_{\mathbf{p}} \mathbf{K} \mathbf{T}_j (\mathbf{T}_i \mathbf{p}_w)^{\odot} \\ \frac{\partial \mathbf{p}^{'}_w}{\partial \boldsymbol{\xi_j}} &= d^{jL}_{\mathbf{p}} \mathbf{K} \frac{\partial (\mathbf{T}_j \mathbf{T}_i^{-1} \mathbf{p}^{'}_w)}{\partial \boldsymbol{\xi_i}} = d^{jL}_{\mathbf{p}} \mathbf{K} (\mathbf{T}_j \mathbf{T}_i^{-1} \mathbf{p}_w)^{\odot} \end{aligned} \qquad (2.2)$$

add detail scalar derivation…….

Secondly, according to

$$(r^d_{\mathbf{p}})_{ij} := I^L_j(\mathbf{p}^{'}) - b^L_j - \frac{e^{a^L_j}}{e^{a^L_i}} (I^L_i(\mathbf{p}) - b^L_i) \qquad (2.2)$$

We have:

$$\begin{aligned} &\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial a_i} = \frac{e^{a^L_j}}{e^{a^L_i}} (I^L_i(\mathbf{p}) - b^L_i), \frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial a_j} = -\frac{e^{a^L_j}}{e^{a^L_i}} (I^L_i(\mathbf{p}) - b^L_i) \\ &\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial b_i} = \frac{e^{a^L_j}}{e^{a^L_i}}, \qquad \frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial b_j} = -1 \end{aligned} \qquad (2.2)$$

add detail Calibration derivation……

$$
\begin{aligned}
&\mathbf{p}^{'} = d_{\mathbf{p}}^{jL}\mathbf{K}(\mathbf{T}_j\mathbf{T}_i^{-1}\mathbf{p}_w) \\
&assume : \mathbf{T}_j\mathbf{T}_i^{-1} = \begin{pmatrix} r_{11}^{ji} & r_{12}^{ji} & r_{13}^{ji} & t_1^{ji} \\ r_{21}^{ji} & r_{22}^{ji} & r_{23}^{ji} & t_2^{ji} \\ r_{31}^{ji} & r_{32}^{ji} & r_{33}^{ji} & t_3^{ji} \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&\mathbf{p}_w^{'} = \begin{pmatrix} r_{11}^{ji} f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) + r_{12}^{ji} f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) + r_{13}^{ji}(d_{\mathbf{p}}^{iL})^{-1} + t_1^{ji} \\ r_{21}^{ji} f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) + r_{22}^{ji} f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) + r_{23}^{ji}(d_{\mathbf{p}}^{iL})^{-1} + t_2^{ji} \\ r_{31}^{ji} f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) + r_{32}^{ji} f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) + r_{33}^{ji}(d_{\mathbf{p}}^{iL})^{-1} + t_3^{ji} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \frac{a}{d_{\mathbf{p}}^{iL}} + t_1^{ji} \\ \frac{b}{d_{\mathbf{p}}^{iL}} + t_2^{ji} \\ \frac{c}{d_{\mathbf{p}}^{iL}} + t_3^{ji} \\ 1 \end{pmatrix} = \begin{pmatrix} x^{'} \\ y^{'} \\ z^{'} \\ 1 \end{pmatrix} \Rightarrow d_{\mathbf{p}}^{jL} = (z^{'})^{-1}, \mathbf{p}^{'} = \begin{pmatrix} f_x x^{'} d_{\mathbf{p}}^{jL} + c_x \\ f_y y^{'} d_{\mathbf{p}}^{jL} + c_y \\ 1 \\ d_{\mathbf{p}}^{jL} \end{pmatrix} \\
&\frac{\partial(I_j^L(\mathbf{p}^{'}))}{\partial \mathbf{p}^{'}} = (g_x^{'}, g_y^{'}, 0, 0) \\
&\frac{\partial \mathbf{p}^{'}}{\partial \mathbf{p}_w^{'}} = \begin{pmatrix} f_x(z^{'})^{-1} & 0 & -x^{'} f_x (z^{'})^{-2} & 0 \\ 0 & f_y(z^{'})^{-1} & -y^{'} f_y (z^{'})^{-2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & (z^{'})^{-2} & 0 \end{pmatrix}, \frac{\partial \mathbf{p}_w^{'}}{\partial d_{\mathbf{p}}^{iL}} = \begin{pmatrix} -\frac{a}{(d_{\mathbf{p}}^{iL})^2} \\ -\frac{b}{(d_{\mathbf{p}}^{iL})^2} \\ -\frac{c}{(d_{\mathbf{p}}^{iL})^2} \\ 0 \end{pmatrix} \\
&\Rightarrow \frac{\partial (r_{\mathbf{p}}^d)_{ij}}{\partial d_{\mathbf{p}}^{iL}} = \frac{\partial(I_j^L(\mathbf{p}^{'}))}{\partial d_{\mathbf{p}}^{iL}} = \frac{\partial(I_j^L(\mathbf{p}^{'}))}{\partial \mathbf{p}^{'}} \frac{\partial \mathbf{p}^{'}}{\partial \mathbf{p}_w^{'}} \frac{\partial \mathbf{p}_w^{'}}{\partial d_{\mathbf{p}}^{iL}} \\
&= -\frac{g_x^{'} f_x a}{z^{'}(d_{\mathbf{p}}^{iL})^2} - \frac{g_y^{'} f_y b}{z^{'}(d_{\mathbf{p}}^{iL})^2} + \frac{c(g_x^{'} x^{'} f_x + g_y^{'} y^{'} f_y)}{(z^{'} d_{\mathbf{p}}^{iL})^2} \\
&= \frac{c(g_x^{'} x^{'} f_x + g_y^{'} y^{'} f_y) - g_x^{'} f_x a z^{'} - g_y^{'} f_y b z^{'}}{(z^{'} d_{\mathbf{p}}^{iL})^2}
\end{aligned} \tag{2.2}
$$

1.2.3.2 Static Parameter

Firstly, For a stereo frame i: inverse depth $d_{\mathbf{p}}^{iL} = d_{\mathbf{p}}^{iR}$, a left frame iL pixel $\mathbf{p}$ is projected to right frame iR with $\mathbf{p}^{'}$:

$$
\begin{aligned}
&\mathbf{p} = \begin{pmatrix} u^i \\ v^i \\ 1 \\ d_{\mathbf{p}}^{iL} \end{pmatrix}, \mathbf{p}_w = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, d_{\mathbf{p}}^{iL} = z^{-1}, \mathbf{p}_w = (d_{\mathbf{p}}^{iL})^{-1}\mathbf{K}^{-1}\mathbf{p} \\
&= \begin{pmatrix} f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) \\ f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) \\ (d_{\mathbf{p}}^{iL})^{-1} \\ 1 \end{pmatrix}, \mathbf{T}_{RL} = \begin{pmatrix} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&\mathbf{p}^{'} = d_{\mathbf{p}}^{iR}\mathbf{K}(\mathbf{T}_{RL}\mathbf{p}_w) \\
&= d_{\mathbf{p}}^{iL} \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) + t_1 \\ f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) \\ (d_{\mathbf{p}}^{iL})^{-1} \\ 1 \end{pmatrix} = \begin{pmatrix} u^i + t_1 f_x d_{\mathbf{p}}^{iL} \\ v^i \\ 1 \\ d_{\mathbf{p}}^{iL} \end{pmatrix} \qquad (1) \\
&\frac{\partial r_{\mathbf{p}}^s}{\partial d_{\mathbf{p}}^{iL}} = \frac{\partial(I_i^R(\mathbf{p}^{'})) - \frac{e^{a_i^R}}{e^{a_i^L}}(I_i^L(\mathbf{p}))}{\partial d_{\mathbf{p}}^{iL}} = \left(\frac{\partial(I_i^R(\mathbf{p}^{'}))}{\partial \mathbf{p}^{'}} - \frac{e^{a_i^R}}{e^{a_i^L}}\frac{\partial(I_i^L(\mathbf{p}))}{\partial \mathbf{p}^{'}}\right)\frac{\partial \mathbf{p}^{'}}{\partial d_{\mathbf{p}}^{iL}} \\
&= [(g_x^{iR}, g_y^{iR}, 0, 0) - \mathbf{0}^T] \begin{pmatrix} t_1 f_x \\ 0 \\ 0 \\ 1 \end{pmatrix} = g_x^{iR} t_1 f_x
\end{aligned}
$$

Secondly, according to:

$$
r_{\mathbf{p}}^s := I_i^R(\mathbf{p}^{'}) - b_i^R - \frac{e^{a_i^R}}{e^{a_i^L}}(I_i^L(\mathbf{p}) - b_i^L) \qquad (2.2)
$$

We have:

$$
\begin{aligned}
&\frac{\partial (r_{\mathbf{p}}^s)_{ij}}{\partial a_i} = \frac{e^{a_j^L}}{e^{a_i^L}}(I_i^L(\mathbf{p}) - b_i^L), \frac{\partial (r_{\mathbf{p}}^s)_{ij}}{\partial a_j} = -\frac{e^{a_j^L}}{e^{a_i^L}}(I_i^L(\mathbf{p}) - b_i^L) \\
&\frac{\partial (r_{\mathbf{p}}^s)_{ij}}{\partial b_i} = \frac{e^{a_j^L}}{e^{a_i^L}}, \qquad \frac{\partial (r_{\mathbf{p}}^s)_{ij}}{\partial b_j} = -1
\end{aligned} \qquad (2.2)
$$

add detail Calibration derivation……

Chapter2 INERTIAL RESIDUALS

2.1 INTRODUCTION

Windowed Optimization is a classic method in non-linear optimization.

2.1.1 NOTATION

Throughout the paper, we will write matrices as bold capital letters ($\mathbf{R}$) and vectors as bold lower case letters ($\boldsymbol{\xi}$), light lower-case letters to denote scalars (s). Light upper-case letters are used to represent functions (I).

Homogeneous camera calibration matrices are denoted by $\mathbf{K}$ as (2.1). Camera poses are represented by matrices of the special Euclidean group $\mathbf{T} \in SE(3)$, which transform a 3D coordinate from the camera coordinate system to the world coordinate system. In this paper, a homogeneous 2D image coordinate point $\mathbf{p}$ is represented by its image coordinate and inverse depth as (2.1) relative to its host keyframe I_i^L. The host keyframe is the frame the point got selected from. Corresponding homogeneous 3D world coordinate point $\mathbf{p}_w$ is denoted as (2.1). $\Pi_{\mathbf{K}}$ are used to denote camera projection functions. The jacobian of I_i, $\Pi_{\mathbf{K}}$ is denoted as (2.1)

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{K}^{-1} = \begin{pmatrix} f_x^{-1} & 0 & -f_x^{-1}c_x & 0 \\ 0 & f_y^{-1} & -f_y^{-1}c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$
$$\mathbf{p} = \begin{pmatrix} u^i \\ v^i \\ 1 \\ d_{\mathbf{p}} \end{pmatrix}, \mathbf{p}_w = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, d_{\mathbf{p}} = z^{-1}, \mathbf{p} = d_{\mathbf{p}}\mathbf{K}\mathbf{p}_w = \Pi_{\mathbf{K}}(\mathbf{p}_w) \tag{2.1}$$
$$\frac{\partial I_i^L(\mathbf{p})}{\partial \mathbf{p}} = (g_x, g_y, 0, 0), \frac{\partial \mathbf{p}}{\partial \mathbf{p}_w} = \begin{pmatrix} f_x z^{-1} & 0 & -x f_x z^{-2} & 0 \\ 0 & f_y z^{-1} & -y f_y z^{-2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & z^{-2} & 0 \end{pmatrix}$$

1.1.2 QUESTION IMPORT

Assume we observe 5 points $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5\}$ in 4 keyframes $\boldsymbol{\mathcal{F}} = \{I_1, I_2, I_3, I_4\}$, every keyframe has stereo vision (I_i^L, I_i^R) abbreviated as (iL, iR). A point can also be observed by other frame as shown in Table(2.1). Question is how to use Windowed Optimization method to make our observation more accurate ?

Table (2.1)

Image point	Host keyframe	Observe by
$\mathbf{p}_1$	$1L$	$1R, 2L$
$\mathbf{p}_2$	$2L$	$2R, 1L, 3L$
$\mathbf{p}_3$	$2L$	$3L$
$\mathbf{p}_4$	$3L$	$1L, 2L$
$\mathbf{p}_5$	$4L$	$3L, 4L$

2.2 SOLUTION

We use direct method to construct residual, Windowed Gauss-Newton method to optimization residual。

1.2.1 CONSTRUCT RESIDUAL

Dynamic multi-view stereo residuals $E_{ij}^{\mathbf{p}}$ are defined as

$$E_{ij}^{\mathbf{p}} = w_{\mathbf{p}}||(r_{\mathbf{p}}^d)_{ij}||_\gamma, \quad (r_{\mathbf{p}}^d)_{ij} := I_j^L(\mathbf{p}') - b_j^L - \frac{e^{a_j^L}}{e^{a_i^L}}(I_i^L(\mathbf{p}) - b_i^L) \tag{2.2}$$

γ is Huber norm. a_i^L, b_i^L is affine brightness parameters to frame iL . $w_{\mathbf{p}}$ is a gradient-dependent weighting parameters, $\mathbf{p}$ in frame I_i^L projected to I_j^L is $\mathbf{p}'$ as:

$$w_{\mathbf{p}} := \frac{c^2}{c^2 + ||\nabla I_i(\mathbf{p})||_2^2}, \quad \mathbf{p}' = d_{\mathbf{p}}^{jL}\mathbf{K}(\mathbf{T}_j\mathbf{T}_i^{-1}((d_{\mathbf{p}}^{iL})^{-1}\mathbf{K}^{-1}\mathbf{p})) \tag{2.3}$$

Static one-view stereo residuals $E_{is}^{\mathbf{p}}$ are modified to

$$E_{is}^{\mathbf{p}} = w_{\mathbf{p}}||r_{\mathbf{p}}^s||_\gamma, \quad r_{\mathbf{p}}^s := I_i^R(\mathbf{p}') - b_i^R - \frac{e^{a_i^R}}{e^{a_i^L}}(I_i^L(\mathbf{p}) - b_i^L) \tag{2.2}$$

Hostframe of $\mathbf{p}$ is I_i^L. a_i^R, b_i^R is affine brightness parameters to frame iR. $\mathbf{p}$ in frame I_i^L projected to I_i^R is $\mathbf{p}'$ as :

$$\mathbf{p}' = d_{\mathbf{p}}^{iR}\mathbf{K}(\mathbf{T}_{RL}((d_{\mathbf{p}}^{iL})^{-1}\mathbf{K}^{-1}\mathbf{p})) \tag{2.2}$$

Total residuals

$$E(\delta) = \sum_{i\in\mathcal{F}}\sum_{\mathbf{p}\in\mathcal{P}_i}\Big(\sum_{j\in obs^t(\mathbf{P})} E_{ij}^{\mathbf{p}} + \lambda E_{is}^{\mathbf{p}}\Big) \tag{2.2}$$

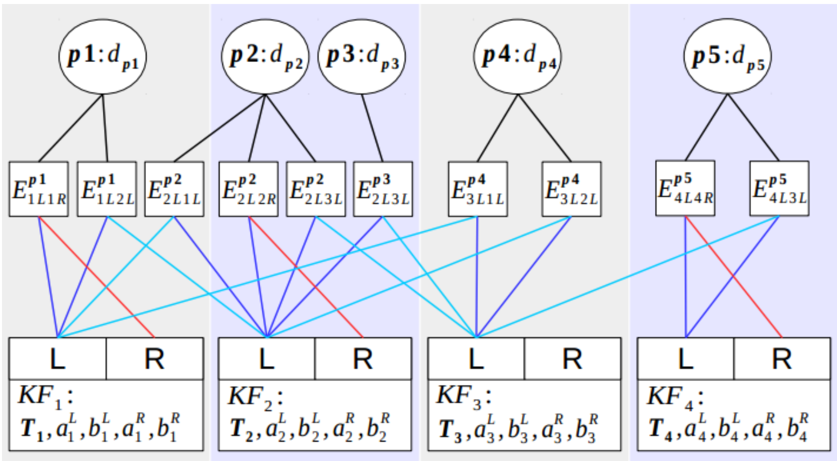
To balance the relative weights of temporal multi-view and static stereo, we introduce a coupling factor λ to weight the constraints from static stereo differently. $\mathcal{P}_i$ is a set of all image point host by frame iL. $obs^t(\mathbf{p})$ are the observations of $\mathbf{p}$ from

$$\boldsymbol{\delta} = \begin{pmatrix} (\boldsymbol{\xi}_1^T, \ldots, \boldsymbol{\xi}_{N_f}^T)^T \\ (d_{\mathbf{p}_1}, \ldots, d_{\mathbf{p}_{N_p}})^T \\ (a_1^L, a_1^R, b_1^L, b_1^R)^T \\ \vdots \\ (a_{N_f}^L, a_{N_f}^R, b_{N_f}^L, b_{N_f}^R)^T \\ (f_x, f_y, c_x, c_y)^T \end{pmatrix} \in \mathbb{R}^{10N_f+N_p+4}, \boldsymbol{\xi}_i = (ln\mathbf{T}_i)^{\vee} \in \mathbb{R}^6 \tag{2.1}$$

temporal multi-view stereo. If there are N_p image point and N_f keyframes in $\boldsymbol{\mathcal{F}}$, optimization variable $\boldsymbol{\delta}$ is

In this example, there are 7 dynamic residuals and 3 static residuals, Factor graph of the residuals function is

Total residuals is



$$\begin{aligned} E(\delta) &= E^{\mathbf{p}_1}_{1L2L} + E^{\mathbf{p}_2}_{2L1L} + E^{\mathbf{p}_2}_{2L3L} + E^{\mathbf{p}_3}_{2L3L} + E^{\mathbf{p}_4}_{3L1L} + E^{\mathbf{p}_4}_{3L2L} + E^{\mathbf{p}_5}_{4L3L} \\ &\quad + E^{\mathbf{p}_1}_{1L1R} + E^{\mathbf{p}_2}_{2L2R} + E^{\mathbf{p}_5}_{4L4R} \\ &= E_d(\delta) + E_s(\delta) \end{aligned}$$

$$E_s(\delta) = \begin{pmatrix} r^s_{\mathbf{p}_1} \\ r^s_{\mathbf{p}_2} \\ r^s_{\mathbf{p}_5} \end{pmatrix}^T \begin{pmatrix} \lambda w_{\mathbf{p}_1} & 0 & 0 \\ 0 & \lambda w_{\mathbf{p}_2} & 0 \\ 0 & 0 & \lambda w_{\mathbf{p}_5} \end{pmatrix} \begin{pmatrix} r^s_{\mathbf{p}_1} \\ r^s_{\mathbf{p}_2} \\ r^s_{\mathbf{p}_5} \end{pmatrix} = (\mathbf{r}^s)^T \mathbf{W}^s \mathbf{r}^s$$

$$\mathbf{J}_s = \begin{pmatrix} \frac{\partial r^s_{\mathbf{p}_1}}{\partial \boldsymbol{\xi}_1} \cdots \frac{\partial r^s_{\mathbf{p}_1}}{\partial \boldsymbol{\xi}_4} & \frac{\partial r^s_{\mathbf{p}_1}}{\partial d_{\mathbf{p}_1}} \cdots \frac{\partial r^s_{\mathbf{p}_1}}{\partial d_{\mathbf{p}_5}} & \frac{\partial r^s_{\mathbf{p}_1}}{\partial a^L_1} \cdots \frac{\partial r^s_{\mathbf{p}_1}}{\partial b^R_4} & \frac{\partial r^s_{\mathbf{p}_1}}{\partial f_x} \cdots \frac{\partial r^s_{\mathbf{p}_1}}{\partial c_y} \\ \frac{\partial r^s_{\mathbf{p}_2}}{\partial \boldsymbol{\xi}_1} \cdots \frac{\partial r^s_{\mathbf{p}_2}}{\partial \boldsymbol{\xi}_4} & \frac{\partial r^s_{\mathbf{p}_2}}{\partial d_{\mathbf{p}_1}} \cdots \frac{\partial r^s_{\mathbf{p}_2}}{\partial d_{\mathbf{p}_5}} & \frac{\partial r^s_{\mathbf{p}_2}}{\partial a^L_1} \cdots \frac{\partial r^s_{\mathbf{p}_2}}{\partial b^R_4} & \frac{\partial r^s_{\mathbf{p}_2}}{\partial f_x} \cdots \frac{\partial r^s_{\mathbf{p}_2}}{\partial c_y} \\ \frac{\partial r^s_{\mathbf{p}_5}}{\partial \boldsymbol{\xi}_1} \cdots \frac{\partial r^s_{\mathbf{p}_5}}{\partial \boldsymbol{\xi}_4} & \frac{\partial r^s_{\mathbf{p}_5}}{\partial d_{\mathbf{p}_1}} \cdots \frac{\partial r^s_{\mathbf{p}_5}}{\partial d_{\mathbf{p}_5}} & \frac{\partial r^s_{\mathbf{p}_5}}{\partial a^L_1} \cdots \frac{\partial r^s_{\mathbf{p}_5}}{\partial b^R_4} & \frac{\partial r^s_{\mathbf{p}_5}}{\partial f_x} \cdots \frac{\partial r^s_{\mathbf{p}_5}}{\partial c_y} \end{pmatrix}_{3\times 49} \tag{2.2}$$

$$E_d(\delta) = \begin{pmatrix} (r^d_{\mathbf{p}_1})_{12} \\ (r^d_{\mathbf{p}_1})_{21} \\ \vdots \\ (r^d_{\mathbf{p}_5})_{43} \end{pmatrix}^T \begin{pmatrix} w_{\mathbf{p}_1} & 0 & \cdots & 0 \\ 0 & w_{\mathbf{p}_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{\mathbf{p}_5} \end{pmatrix} \begin{pmatrix} (r^d_{\mathbf{p}_1})_{12} \\ (r^d_{\mathbf{p}_1})_{21} \\ \vdots \\ (r^d_{\mathbf{p}_5})_{43} \end{pmatrix} = (\mathbf{r}^d)^T \mathbf{W}^d \mathbf{r}^d$$

$$
\mathbf{J}_d = \begin{pmatrix}
\frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial \boldsymbol{\xi}_1} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial \boldsymbol{\xi}_4} & \frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial d_{\mathbf{p_1}}} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial d_{\mathbf{p_5}}} & \frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial a^L_1} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial b^R_4} & \frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial f_x} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{12}}{\partial c_y} \\
\frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial \boldsymbol{\xi}_1} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial \boldsymbol{\xi}_4} & \frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial d_{\mathbf{p_1}}} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial d_{\mathbf{p_5}}} & \frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial a^L_1} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial b^R_4} & \frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial f_x} \cdots & \frac{\partial (r^d_{\mathbf{p_1}})_{21}}{\partial c_y} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial \boldsymbol{\xi}_1} \cdots & \frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial \boldsymbol{\xi}_4} & \frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial d_{\mathbf{p_1}}} \cdots & \frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial d_{\mathbf{p_5}}} & \frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial a^L_1} \cdots & \frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial b^R_4} & \frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial f_x} \cdots & \frac{\partial (r^d_{\mathbf{p_5}})_{43}}{\partial c_y}
\end{pmatrix}_{7\times 49} \tag{2.2}
$$

Iteration δ^* can be calculated by

$$
\begin{gathered}
(\mathbf{J}_s^T \lambda \mathbf{W}^s \mathbf{J}_s + \mathbf{J}_d^T \mathbf{W}^d \mathbf{J}_d)\delta^* = -(\mathbf{J}_s^T \lambda \mathbf{W}^s \mathbf{r}^s + \mathbf{J}_d^T \mathbf{W}^d \mathbf{r}^d) \\
\mathbf{J}_s \in \mathbb{R}^{3\times 49}, \mathbf{W}^s \in \mathbb{R}^{3\times 3}, \mathbf{J}_s \in \mathbb{R}^{7\times 49}, \mathbf{W}^s \in \mathbb{R}^{7\times 7},
\end{gathered} \tag{2.2}
$$

We construct residuals and its formulation.

1.2.2 JACOBIAN CITATION

We know for a Lie algebra $\boldsymbol{\rho} \in \mathbb{R}^3, \boldsymbol{\phi} \in \mathbb{R}^3, \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{pmatrix} \in \mathbb{R}^6$ and $\mathbf{p}_w$:

$$
\begin{aligned}
&\boldsymbol{\xi}^{\wedge} = \begin{pmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{pmatrix}^{\wedge} = \begin{pmatrix} \boldsymbol{\phi}^{\wedge} & \boldsymbol{\rho} \\ \mathbf{0}^T & \mathbf{0}^T \end{pmatrix} \in \mathbb{R}^{4\times 4} \\
&\boldsymbol{\epsilon} \in \mathbb{R}^3, \begin{pmatrix} \boldsymbol{\epsilon} \\ 1 \end{pmatrix}^{\odot} = \begin{pmatrix} \mathbf{E} & -\boldsymbol{\epsilon}^{\wedge} \\ \mathbf{0}^T & \mathbf{0}^T \end{pmatrix} \in \mathbb{R}^{4\times 6} \\
&\frac{\partial (exp(\xi^{\wedge})\mathbf{p}_w)}{\partial \boldsymbol{\xi}} = \frac{\partial (\mathbf{T}\mathbf{p}_w)}{\partial \boldsymbol{\xi}} = (\mathbf{T}\mathbf{p}_w)^{\odot} \\
&\mathbf{T}\mathbf{p}_w = exp(\xi^{\wedge})\mathbf{p}_w \approx (\mathbf{E} + \xi^{\wedge})\mathbf{p}_w \\
&\frac{\partial (exp(\xi^{\wedge})\mathbf{p}_w)}{\partial \boldsymbol{\xi}} \approx \frac{\partial (\mathbf{E} + \xi^{\wedge})}{\partial \boldsymbol{\xi}} = \mathbf{0} + \frac{\partial (\xi^{\wedge}\mathbf{p}_w)}{\partial \boldsymbol{\xi}} \approx (\mathbf{T}\mathbf{p}_w)^{\odot} \\
&since, \frac{\partial (\mathbf{T}\mathbf{p}_w)}{\partial \boldsymbol{\xi}} = (\mathbf{T}^{-1}\mathbf{p}_w)^{\odot} = \frac{\partial (exp(-\xi^{\wedge})\mathbf{p}_w)}{\partial \boldsymbol{\xi}} \\
&= \frac{\partial (\mathbf{E} - \xi^{\wedge})}{\partial \boldsymbol{\xi}} = -(\mathbf{T}\mathbf{p}_w)^{\odot}
\end{aligned} \tag{2.2}
$$

1.2.3 JACOBIAN DERIVATION

1.2.3.1 Dynamic Parameter

Firstly, if $\mathbf{p}$ is neither observed by frame mL, mR nor hosted by nL, nR:

$$
\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial \boldsymbol{\xi_m}} = \frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial \boldsymbol{\xi_n}} = \mathbf{0}^T, so \quad \frac{\partial (r^d_{\mathbf{p}_1})_{12}}{\partial \boldsymbol{\xi}_3} = \frac{\partial r^d_{(\mathbf{p}_1})_{12}}{\partial \boldsymbol{\xi}_4} = \ldots = \mathbf{0}^T, \tag{2.2}
$$

otherwise, we follow

$$\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial \boldsymbol{\xi_i}} = \frac{\partial (I^L_j(\mathbf{p}'))}{\partial \boldsymbol{\xi_i}} = \frac{\partial (I^L_j(\mathbf{p}'))}{\partial \mathbf{p}'}\frac{\partial \mathbf{p}'}{\partial \mathbf{p}'_w}\frac{\partial \mathbf{p}'_w}{\partial \boldsymbol{\xi_i}} \qquad (2.2)$$
$$\mathbf{p}'_w = \mathbf{T}_j\mathbf{T}_i^{-1}\mathbf{p}_w = \mathbf{T}_j\mathbf{T}_i^{-1}((d^{iL}_{\mathbf{p}})^{-1}\mathbf{K}^{-1}\mathbf{p})$$

For one frame iL, we have $\mathbf{p}$ and $\mathbf{K}$, then we can get

$$\left\{ \mathbf{p}_w = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} f_x^{-1}(d^{iL}_{\mathbf{p}})^{-1}(u^i - c_x) \\ f_y^{-1}(d^{iL}_{\mathbf{p}})^{-1}(v^i - c_y) \\ (d^{iL}_{\mathbf{p}})^{-1} \\ 1 \end{pmatrix} \right. \qquad (2.2)$$

$$\frac{\partial \mathbf{p}'_w}{\partial \boldsymbol{\xi_i}} = d^{jL}_{\mathbf{p}}\mathbf{K}\mathbf{T}_j\frac{\partial(\mathbf{T}_i^{-1}\mathbf{p}'_w)}{\partial \boldsymbol{\xi_i}} = -d^{jL}_{\mathbf{p}}\mathbf{K}\mathbf{T}_j(\mathbf{T}_i\mathbf{p}_w)^{\odot}$$
$$\frac{\partial \mathbf{p}'_w}{\partial \boldsymbol{\xi_j}} = d^{jL}_{\mathbf{p}}\mathbf{K}\frac{\partial(\mathbf{T}_j\mathbf{T}_i^{-1}\mathbf{p}'_w)}{\partial \boldsymbol{\xi_i}} = d^{jL}_{\mathbf{p}}\mathbf{K}(\mathbf{T}_j\mathbf{T}_i^{-1}\mathbf{p}_w)^{\odot} \qquad (2.2)$$

add detail scalar derivation…….

Secondly, according to

$$(r^d_{\mathbf{p}})_{ij} := I^L_j(\mathbf{p}') - b^L_j - \frac{e^{a^L_j}}{e^{a^L_i}}(I^L_i(\mathbf{p}) - b^L_i) \qquad (2.2)$$

We have:

$$\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial a_i} = \frac{e^{a^L_j}}{e^{a^L_i}}(I^L_i(\mathbf{p}) - b^L_i), \frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial a_j} = -\frac{e^{a^L_j}}{e^{a^L_i}}(I^L_i(\mathbf{p}) - b^L_i)$$
$$\frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial b_i} = \frac{e^{a^L_j}}{e^{a^L_i}}, \qquad \frac{\partial (r^d_{\mathbf{p}})_{ij}}{\partial b_j} = -1 \qquad (2.2)$$

add detail Calibration derivation……

$$
\begin{aligned}
&\mathbf{p}^{'} = d_{\mathbf{p}}^{jL}\mathbf{K}(\mathbf{T}_j\mathbf{T}_i^{-1}\mathbf{p}_w)\\
&assume: \mathbf{T}_j\mathbf{T}_i^{-1} = \begin{pmatrix} r_{11}^{ji} & r_{12}^{ji} & r_{13}^{ji} & t_1^{ji} \\ r_{21}^{ji} & r_{22}^{ji} & r_{23}^{ji} & t_2^{ji} \\ r_{31}^{ji} & r_{32}^{ji} & r_{33}^{ji} & t_3^{ji} \\ 0 & 0 & 0 & 1 \end{pmatrix}\\
&\mathbf{p}_w^{'} = \begin{pmatrix} r_{11}^{ji}f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) + r_{12}^{ji}f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) + r_{13}^{ji}(d_{\mathbf{p}}^{iL})^{-1} + t_1^{ji} \\ r_{21}^{ji}f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) + r_{22}^{ji}f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) + r_{23}^{ji}(d_{\mathbf{p}}^{iL})^{-1} + t_2^{ji} \\ r_{31}^{ji}f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i - c_x) + r_{32}^{ji}f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i - c_y) + r_{33}^{ji}(d_{\mathbf{p}}^{iL})^{-1} + t_3^{ji} \\ 1 \end{pmatrix}\\
&= \begin{pmatrix} \frac{a}{d_{\mathbf{p}}^{iL}} + t_1^{ji} \\ \frac{b}{d_{\mathbf{p}}^{iL}} + t_2^{ji} \\ \frac{c}{d_{\mathbf{p}}^{iL}} + t_3^{ji} \\ 1 \end{pmatrix} = \begin{pmatrix} x^{'} \\ y^{'} \\ z^{'} \\ 1 \end{pmatrix} \Rightarrow d_{\mathbf{p}}^{jL} = (z^{'})^{-1}, \mathbf{p}^{'} = \begin{pmatrix} f_x x^{'} d_{\mathbf{p}}^{jL} + c_x \\ f_y y^{'} d_{\mathbf{p}}^{jL} + c_y \\ 1 \\ d_{\mathbf{p}}^{jL} \end{pmatrix}\\
&\frac{\partial(I_j^L(\mathbf{p}^{'}))}{\partial\mathbf{p}^{'}} = (g_x^{'}, g_y^{'}, 0, 0)\\
&\frac{\partial\mathbf{p}^{'}}{\partial\mathbf{p}_w^{'}} = \begin{pmatrix} f_x(z^{'})^{-1} & 0 & -x^{'}f_x(z^{'})^{-2} & 0 \\ 0 & f_y(z^{'})^{-1} & -y^{'}f_y(z^{'})^{-2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & (z^{'})^{-2} & 0 \end{pmatrix}, \frac{\partial\mathbf{p}_w^{'}}{\partial d_{\mathbf{p}}^{iL}} = \begin{pmatrix} -\frac{a}{(d_{\mathbf{p}}^{iL})^2} \\ -\frac{b}{(d_{\mathbf{p}}^{iL})^2} \\ -\frac{c}{(d_{\mathbf{p}}^{iL})^2} \\ 0 \end{pmatrix}\\
&\Rightarrow \frac{\partial(r_{\mathbf{p}}^d)_{ij}}{\partial d_{\mathbf{p}}^{iL}} = \frac{\partial(I_j^L(\mathbf{p}^{'}))}{\partial d_{\mathbf{p}}^{iL}} = \frac{\partial(I_j^L(\mathbf{p}^{'}))}{\partial\mathbf{p}^{'}}\frac{\partial\mathbf{p}^{'}}{\partial\mathbf{p}_w^{'}}\frac{\partial\mathbf{p}_w^{'}}{\partial d_{\mathbf{p}}^{iL}}\\
&= -\frac{g_x^{'}f_x a}{z^{'}(d_{\mathbf{p}}^{iL})^2} - \frac{g_y^{'}f_y b}{z^{'}(d_{\mathbf{p}}^{iL})^2} + \frac{c(g_x^{'}x^{'}f_x + g_y^{'}y^{'}f_y)}{(z^{'}d_{\mathbf{p}}^{iL})^2}\\
&= \frac{c(g_x^{'}x^{'}f_x + g_y^{'}y^{'}f_y) - g_x^{'}f_x a z^{'} - g_y^{'}f_y b z^{'}}{(z^{'}d_{\mathbf{p}}^{iL})^2}
\end{aligned}
\tag{2.2}
$$

1.2.3.2 Static Parameter

Firstly, For a stereo frame i: inverse depth $d_{\mathbf{p}}^{iL}=d_{\mathbf{p}}^{iR}$, a left frame iL pixel $\mathbf{p}$ is projected to right frame iR with $\mathbf{p}^{'}$:

$$
\begin{aligned}
&\mathbf{p}=\begin{pmatrix}u^i\\v^i\\1\\d_{\mathbf{p}}^{iL}\end{pmatrix},\mathbf{p}_w=\begin{pmatrix}x\\y\\z\\1\end{pmatrix},d_{\mathbf{p}}^{iL}=z^{-1},\mathbf{p}_w=(d_{\mathbf{p}}^{iL})^{-1}\mathbf{K}^{-1}\mathbf{p}\\
&=\begin{pmatrix}f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i-c_x)\\f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i-c_y)\\(d_{\mathbf{p}}^{iL})^{-1}\\1\end{pmatrix},\mathbf{T}_{RL}=\begin{pmatrix}1&0&0&t_1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\\
&\mathbf{p}^{'}=d_{\mathbf{p}}^{iR}\mathbf{K}(\mathbf{T}_{RL}\mathbf{p}_w)\\
&=d_{\mathbf{p}}^{iL}\begin{pmatrix}f_x&0&c_x&0\\0&f_y&c_y&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}f_x^{-1}(d_{\mathbf{p}}^{iL})^{-1}(u^i-c_x)+t_1\\f_y^{-1}(d_{\mathbf{p}}^{iL})^{-1}(v^i-c_y)\\(d_{\mathbf{p}}^{iL})^{-1}\\1\end{pmatrix}=\begin{pmatrix}u^i+t_1f_xd_{\mathbf{p}}^{iL}\\v^i\\1\\d_{\mathbf{p}}^{iL}\end{pmatrix}\\
&\frac{\partial r_{\mathbf{p}}^s}{\partial d_{\mathbf{p}}^{iL}}=\frac{\partial(I_i^R(\mathbf{p}^{'}))-\frac{e^{a_i^R}}{e^{a_i^L}}(I_i^L(\mathbf{p}))}{\partial d_{\mathbf{p}}^{iL}}=(\frac{\partial(I_i^R(\mathbf{p}^{'}))}{\partial\mathbf{p}^{'}}-\frac{e^{a_i^R}}{e^{a_i^L}}\frac{\partial(I_i^L(\mathbf{p}))}{\partial\mathbf{p}^{'}})\frac{\partial\mathbf{p}^{'}}{\partial d_{\mathbf{p}}^{iL}}\\
&=[(g_x^{iR},g_y^{iR},0,0)-\mathbf{0}^T]\begin{pmatrix}t_1f_x\\0\\0\\1\end{pmatrix}=g_x^{iR}t_1f_x
\end{aligned}
\tag{1}
$$

Secondly, according to:

$$
r_{\mathbf{p}}^s:=I_i^R(\mathbf{p}^{'})-b_i^R-\frac{e^{a_i^R}}{e^{a_i^L}}(I_i^L(\mathbf{p})-b_i^L) \tag{2.2}
$$

We have:

$$
\begin{aligned}
&\frac{\partial(r_{\mathbf{p}}^s)_{ij}}{\partial a_i}=\frac{e^{a_j^L}}{e^{a_i^L}}(I_i^L(\mathbf{p})-b_i^L),\frac{\partial(r_{\mathbf{p}}^s)_{ij}}{\partial a_j}=-\frac{e^{a_j^L}}{e^{a_i^L}}(I_i^L(\mathbf{p})-b_i^L)\\
&\frac{\partial(r_{\mathbf{p}}^s)_{ij}}{\partial b_i}=\frac{e^{a_j^L}}{e^{a_i^L}},\qquad\frac{\partial(r_{\mathbf{p}}^s)_{ij}}{\partial b_j}=-1
\end{aligned}
\tag{2.2}
$$