**Supplementary Material to:**

**Direct Sparse Visual-Inertial Odometry with Stereo Cameras**

Ziqiang Wang

March, 2019

Ziqiang Wang is an engineer of UISEE Technologies Shanghai Co., Ltd and was a postgraduate student with the college of Electronics and Information Engineering of Tongji Universty, Shanghai, China.(e-mail: ziqiang.wang@ uisee.com).

Content

[Chapter1 Visual-inertial Preliminaries 3](#_Toc3320709)

[Chapter2 IMU Error Factors 4](#_Toc3320710)

[1.1 Time-closest measurements selection strategy 4](#_Toc3320711)

[1.2 Errors and covariance calculation pseudo code 6](#_Toc3320712)

[1.3 Jacobian derivation 6](#_Toc3320713)

[Chapter3 Photo Error Factors 7](#_Toc3320714)

[3.1 Construction residual errors 7](#_Toc3320715)

[3.2 Jacobian citation 9](#_Toc3320716)

[1.3 Jacobian derivation 9](#_Toc3320717)

[1.3.1 Dynamic Parameter 9](#_Toc3320718)

[1.3.2 Static Parameter 12](#_Toc3320719)

# Chapter1 Visual-inertial Preliminaries

In our main paper [Ⅳ], The term  is the right Jacobian of  can be calculated by (1.1).

Homogeneous camera calibration matrices are denoted by  as (1.2.1). and homogeneous 2D image coordinate point  is represented by its image coordinate and inverse depth as (1.2.3) relative to its host keyframe . Corresponding homogeneous 3D camera coordinate point  is denoted as (1.2.4).  are used to denote camera projection functions. The jacobian of ,  is denoted as (1.5)

# Chapter2 IMU Error Factors

## 1.1 Time-closest measurements selection strategy

Dynamic multi-view stereo residuals  are defined as

 is Huber norm.  is affine brightness parameters to frame  .  is a gradient-dependent weighting parameters,  in frame  projected to  is  as:

Static one-view stereo residuals  are modified to

Hostframe of  is .  is affine brightness parameters to frame .  in frame  projected to  is  as :

Total residuals

To balance the relative weights of temporal multi-view and static stereo, we introduce a coupling factor  to weight the constraints from static stereo differently.  is a set of all image point host by frame .  are the observations of  from temporal multi-view stereo. If there are  image point and  keyframes in , optimization variable  is

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

Total residuals is



Iteration  can be calculated by

We construct residuals and its formulation.

## 1.2 Errors and covariance calculation pseudo code

 We know for a Lie algebra  and :

## 1.3 Jacobian derivation

# Chapter3 Photo Error Factors

## 3.1 Construction residual errors

Dynamic multi-view stereo residuals  are defined as

 is Huber norm.  is affine brightness parameters to frame  .  is a gradient-dependent weighting parameters,  in frame  projected to  is  as:

Static one-view stereo residuals  are modified to

Hostframe of  is .  is affine brightness parameters to frame .  in frame  projected to  is  as :

Total residuals

To balance the relative weights of temporal multi-view and static stereo, we introduce a coupling factor  to weight the constraints from static stereo differently.  is a set of all image point host by frame .  are the observations of  from temporal multi-view stereo. If there are  image point and  keyframes in , optimization variable  is

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

Total residuals is



Iteration  can be calculated by

We construct residuals and its formulation.

## 3.2 Jacobian citation

 We know for a Lie algebra  and :

## 1.3 Jacobian derivation

### 1.3.1 Dynamic Parameter

Firstly, if  is neither observed by frame ,  nor hosted by , :

otherwise, we follow

For one frame , we have  and , then we can get

Secondly, according to

We have:

add detail Calibration derivation……



### 1.3.2 Static Parameter

Firstly, For a stereo frame : inverse depth , a left frame  pixel  is projected to right frame  with :



Secondly, according to:

We have:

add detail Calibration derivation……