Parallel Tracking and Mapping for Small AR Workspaces

郭尧

2015.9.6

Introduction

- 一种应用于小场景AR的**单目**跟踪与地图构建方法
- 单目摄像头SLAM的问题

  1.实时性差

  2.误差累积得不到校正
- 与SLAM不同，跟踪与地图构建运行各自独立的线程，地图不要求实时更新，避免了实时更新带来的计算负担，因此可以构建包含更多更密集特征的地图，

Method overview

- 跟踪与地图构建相互独立
- 基于Keyframes构建地图，基于光束法平差
- 地图基于最初的两幅关键帧进行初始化（**5点算法**）
- 基于极线搜索发现更多的地图点

Tracking

- 假设地图已经建立，实时估计摄像头相对于地图的位置
- two-stage tracking procedure

1. A new frame is acquired from the camera, and a prior pose estimate is generated from a motion model.
2. Map points are projected into the image according to the frame's prior pose estimate.
3. A small number (50) of the coarsest-scale features are searched for in the image.
4. The camera pose is updated from these coarse matches.
5. A larger number (1000) of points is re-projected and searched for in the image.
6. A final pose estimate for the frame is computed from all the matches found.

Tracking

- 假设地图已经建立，实时估计摄像头相对于地图的位置
- 具体工作流程

Image acquisition	Camera pose and projection	Patch Search	Pose update

广角摄像头
采样间隔30ms
FAST角点检测

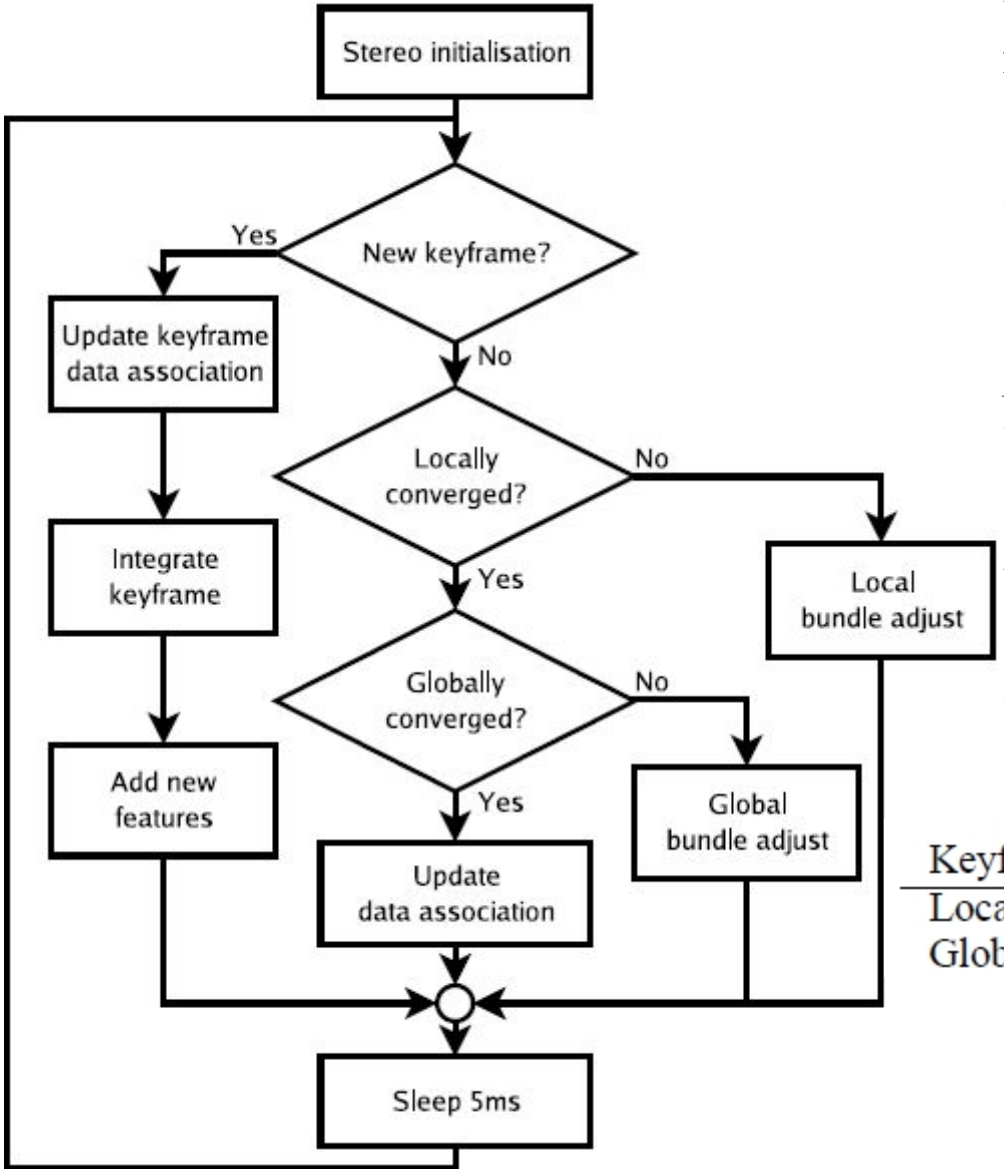
$$p_{jC} = E_{CW} p_{jW} \tag{1}$$

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} = \mathrm{CamProj}(E_{CW} p_{iW}) \tag{2}$$

$$\mathrm{CamProj}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \begin{bmatrix} f_u & 0 \\ 0 & f_v \end{bmatrix} \frac{\mathrm{r}'}{\mathrm{r}} \begin{pmatrix} \frac{x}{z} \\ \frac{y}{z} \end{pmatrix} \tag{3}$$

$$r = \sqrt{\frac{x^2 + y^2}{z^2}} \tag{4}$$

Mapping



Keyframes:

1.Tracking quality must be good;

2.Time since the last keyframe was added must exceed twenty frames;

3.The camera must be a minimum distance away from the nearest keypoint already in the map;

New Map Points:

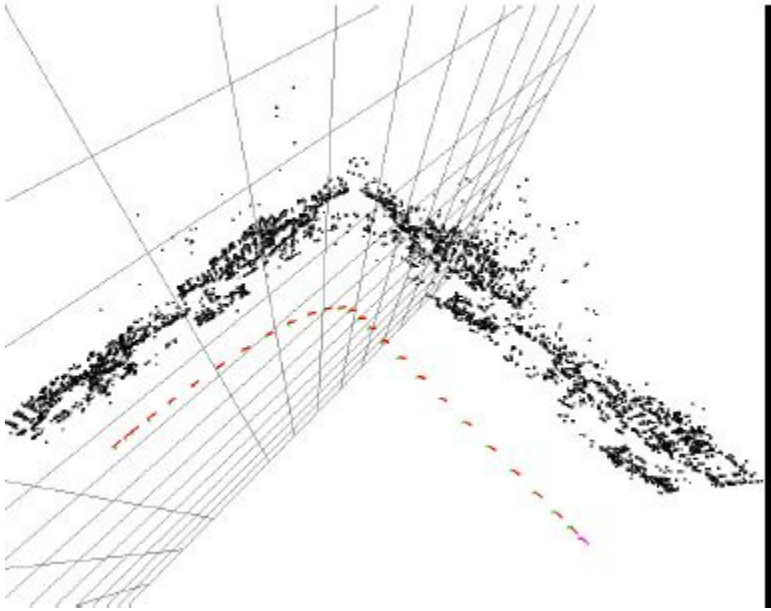
1.Need depth information

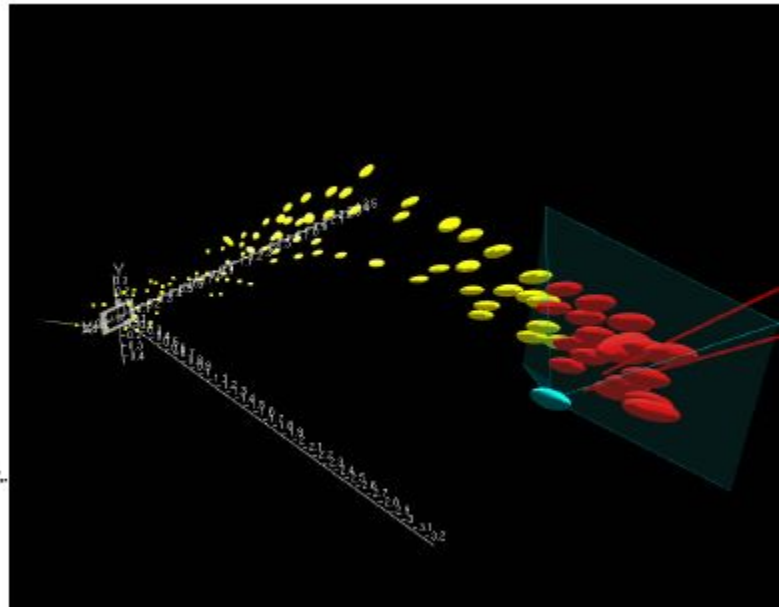
2.Epipolar search between cloest frames

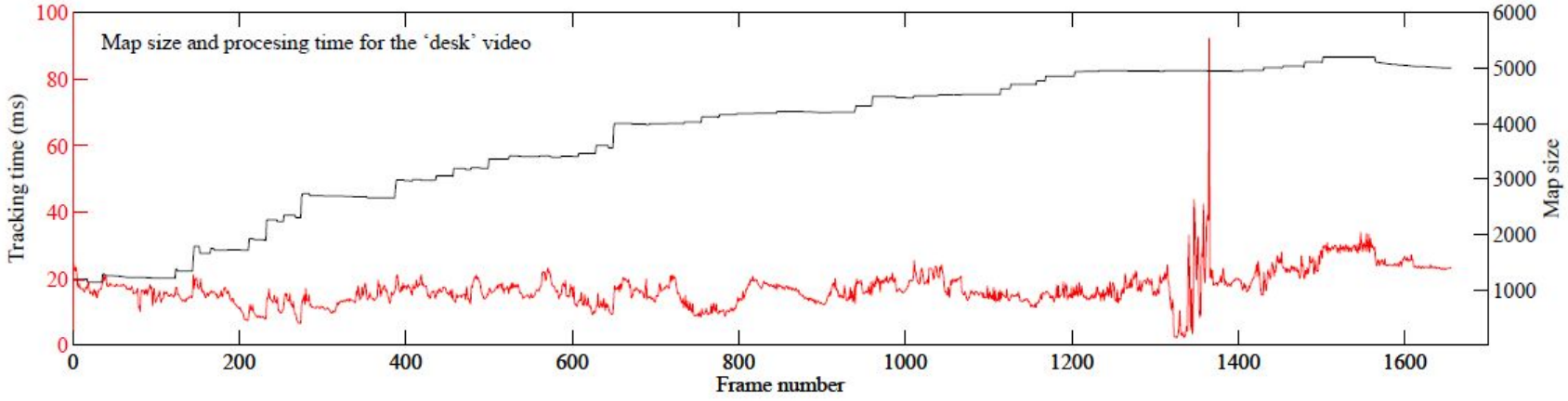
3.Use a prior hypothesis on the likely depth of new candidate points

Keyframes	2-49	50-99	100-149
Local Bundle Adjustment	170ms	270ms	440ms
Global Bundle Adjustment	380ms	1.7s	6.9s

Result






Map size and procesing time for the 'desk' video
Tracking time (ms)
Map size
Frame number
0
20
40
60
80
100
1000
2000
3000
4000
5000
6000
0
200
400
600
800
1000
1200
1400
1600

Problems

Need powerful computing hardware.

Depend on the FAST corner detector.

Limited to repeated structure and lighting changes.It will fail if the real-world scene is substantially and permanently changed