



## (12) 发明专利

(10) 授权公告号 CN 108352071 B

(45) 授权公告日 2022. 05. 13

(21) 申请号 201680061527.2

(22) 申请日 2016.12.29

(65) 同一申请的已公布的文献号  
申请公布号 CN 108352071 A

(43) 申请公布日 2018.07.31

(30) 优先权数据  
7029/CHE/2015 2015.12.29 IN  
15/197,749 2016.06.29 US

(85) PCT国际申请进入国家阶段日  
2018.04.20

(86) PCT国际申请的申请数据  
PCT/US2016/069142 2016.12.29

(87) PCT国际申请的公布数据  
W02017/127220 EN 2017.07.27

(73) 专利权人 德州仪器公司  
地址 美国德克萨斯州

(72) 发明人 索耶·纳戈尔里 马努·马修  
普拉什汉特·拉马纳特普尔·维斯  
瓦纳特  
迪帕克·库马尔·波达尔

(74) 专利代理机构 北京律盟知识产权代理有限公司 11287  
专利代理师 林斯凯

(51) Int.Cl.  
G06T 7/246 (2006.01)  
G01C 11/04 (2006.01)  
G06T 17/00 (2006.01)

(56) 对比文件  
US 2007253618 A1, 2007.11.01  
US 2014139635 A1, 2014.05.22  
US 2014016857 A1, 2014.01.16  
US 2014010407 A1, 2014.01.09

审查员 宋晓毓

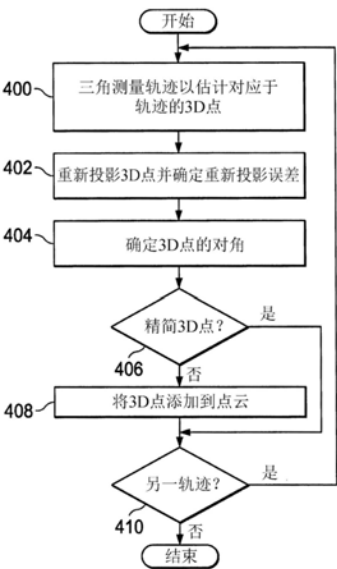
权利要求书6页 说明书7页 附图6页

### (54) 发明名称

用于计算机视觉系统中的运动结构处理的方法

### (57) 摘要

在所描述的实例中,一种用于计算机视觉系统中的运动结构处理的方法包含:接收由单目相机捕获的多个连续帧中的点的所估计二维2D位置的多个轨迹;且由(针对)所述多个轨迹中的每个轨迹(410)产生(408)三维3D点的云;基于所述轨迹中的所述2D位置三角测量(400) 3D点的所估计位置;以及当(406)所述3D点的对角(404)的大小以及所述3D点的投影误差(402)指示所述3D点足够稳固用于包含时,将所述3D点包含于所述云中。



1. 一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中接收由单目相机捕获的多个连续帧中的点的所估计二维2D位置的多个轨迹;以及

针对所述多个轨迹中的每个轨迹,通过以下操作在所述计算机视觉系统中产生三维3D点的云,

基于所述轨迹中的所述2D位置来三角测量3D点的所估计位置;以及

当所述3D点的对角的大小以及所述3D点的投影误差指示所述3D点足够稳固用于包含时,将所述3D点包含于所述云中,其中所述对角的所述大小以及所述投影误差与多对渐小相应阈值中的一或多个进行测试,以确定所述3D点是否足够稳固,其中一对阈值由角大小阈值及投影误差阈值组成。

2. 根据权利要求1所述的方法,其中所述多对渐小相应阈值中的初始阈值对包含基准角大小阈值以及基准投影误差阈值,且通过将对应的预定渐大缩放因子应用于所述基准角大小阈值以及所述基准投影误差阈值来确定所述多对渐小相应阈值中的其它阈值对的阈值。

3. 根据权利要求2所述的方法,其中有五对渐小相应阈值,且所述渐大缩放因子为2、4、8以及16。

4. 根据权利要求1所述的方法,其中N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计所述多个轨迹中的所述2D位置,且每个轨迹包含至多N个2D位置,其中所述多个连续帧中的帧的数目为N,且 $N > 2$ 。

5. 根据权利要求4所述的方法,其中 $N = 6$ 。

6. 根据权利要求1所述的方法,其进一步包括:

将所述云与3D点的先前云合并以产生3D点的最终云,其中,针对所述先前云中的每个3D点,确定是否在所述最终云中包含所述3D点包括:

在所述多个连续帧中的最近帧与所述多个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所述次最近帧中;

如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点包含于所述最终云中。

7. 根据权利要求1所述的方法,其中所述点为关注点。

8. 根据权利要求1所述的方法,其中所述计算机视觉系统包含于汽车安全系统中。

9. 根据权利要求1所述的方法,其进一步包括:

将所述云与3D点的先前云合并以产生3D点的最终云,其中,针对所述先前云中的每个3D点,确定是否在所述最终云中包含所述3D点包括:

在所述多个连续帧中的最近帧与所述多个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所述次最近帧中;

如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以

及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点不包含于所述最终云中。

10. 一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中基于由单目相机捕获的多个连续帧产生三维3D点的云;以及在所述计算机视觉系统中将3D点的所述云与3D点的先前云合并以产生3D点的最终云,其中,针对所述先前云中的每个3D点,确定是否在所述最终云中包含所述3D点包括:

在所述多个连续帧中的最近帧与所述多个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所述次最近帧中;

如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点包含于所述最终云中。

11. 根据权利要求10所述的方法,其中产生3D点的所述云包括:

接收所述多个连续帧中的点的所估计二维2D位置的多个轨迹;以及

通过估计对应于所述多个轨迹中的每个轨迹的3D点来计算3D点的所述云,

其中N个连续帧的滑动窗口中的两个最近帧用于点对应以估计所述轨迹中的所述2D位置,且每个轨迹包含至多N个2D位置,其中所述多个帧中的帧的数目为N,且 $N \geq 2$ 。

12. 根据权利要求11所述的方法,其中 $N=6$ 。

13. 根据权利要求10所述的方法,其中产生3D点的点云包括:

接收所述多个连续帧中的点的所估计二维2D位置的多个轨迹;以及

针对所述多个轨迹中的每个轨迹,通过以下操作产生三维3D点的云,

基于所述轨迹中的所述2D位置来三角测量三维3D点的所估计位置;以及

当所述3D点的对角的大小以及所述3D点的投影误差指示所述3D点足够稳固用于包含时,将所述3D点包含于所述云中,其中所述对角的所述大小以及所述投影误差与多对渐小相应阈值中的一或多个进行测试,以确定所述3D点是否足够稳固,其中一对阈值由角大小阈值及投影误差阈值组成。

14. 根据权利要求13所述的方法,其中所述多对渐小相应阈值中的初始阈值对包含基准角大小阈值以及基准投影误差阈值,且通过将对应的预定渐大缩放因子应用于所述基准角大小阈值以及所述基准投影误差阈值来确定所述多对渐小相应阈值中的其它阈值对的阈值。

15. 根据权利要求14所述的方法,其中有五对渐小相应阈值,且所述渐大缩放因子为2、4、8以及16。

16. 一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中接收由单目相机捕获的N个连续帧中的点的所估计二维2D位置的多个轨迹;

通过估计对应于所述多个轨迹中的每个轨迹的3D点而在所述计算机视觉系统中计算三维3D点云,其中N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计所述轨迹中的

所述2D位置,且每个轨迹包含至多N个2D位置,其中 $N>2$ ;以及

将所述3D点云与先前的3D点云合并以产生最终3D点云,其中,针对所述先前的3D点云中的每个3D点,确定是否在所述最终3D点云中包含所述3D点包括:

在所述N个连续帧中的最近帧与所述N个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所述次最近帧中;

如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点包含于所述最终3D点云中。

17.一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中接收由单目相机捕获的N个连续帧中的点的所估计二维2D位置的多个轨迹;

通过估计对应于所述多个轨迹中的每个轨迹的3D点而在所述计算机视觉系统中计算三维3D点云,其中N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计所述轨迹中的所述2D位置,且每个轨迹包含至多N个2D位置,其中 $N>2$ ;以及

其中计算所述三维3D点云包括:

针对所述多个轨迹中的每个轨迹,通过以下操作产生所述3D点云;

基于所述轨迹中的所述2D位置来三角测量三维3D点的所估计位置;以及

当所述3D点的对角的大小以及所述3D点的投影误差指示所述3D点足够稳固用于包含时,将所述3D点包含于所述3D点云中,其中所述对角的所述大小以及所述投影误差与多对渐小相应阈值中的一或多个进行测试,以确定所述3D点是否足够稳固,其中一对阈值由角大小阈值及投影误差阈值组成。

18.根据权利要求17所述的方法,其中所述多对渐小相应阈值中的初始阈值对包含基准角大小阈值以及基准投影误差阈值,且通过将对应的预定渐大缩放因子应用于所述基准角大小阈值以及所述基准投影误差阈值来确定所述多对渐小相应阈值中的其它阈值对的阈值。

19.根据权利要求18所述的方法,其中有五对渐小相应阈值,且所述渐大缩放因子为2、4、8以及16。

20.一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中接收由单目相机捕获的N个连续帧中的点的所估计二维2D位置的多个轨迹;

通过估计对应于所述多个轨迹中的每个轨迹的3D点而在所述计算机视觉系统中计算三维3D点云,其中N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计所述轨迹中的所述2D位置,且每个轨迹包含至多N个2D位置,其中 $N>2$ ;以及

将所述3D点云与先前的3D点云合并以产生最终3D点云,其中,针对所述先前的3D点云中的每个3D点,确定是否在所述最终3D点云中包含所述3D点包括:

在所述N个连续帧中的最近帧与所述N个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所

述次最近帧中；

如果定位所述第一2D点，那么计算所述第一2D点的极线以及所述第二2D点的极线；以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点包含于所述最终3D点云中。

21. 一种方法，所述方法包括：

在计算机视觉系统中产生由单目相机捕获的多个连续帧；

在所述计算机视觉系统中产生由单目相机捕获的所述多个连续帧中的点的所估计二维2D位置的多个轨迹；以及

针对所述多个轨迹中的每个轨迹，通过以下操作在所述计算机视觉系统中产生三维3D点的云；

基于所述轨迹中的所述2D位置来三角测量3D点的所估计位置；以及

当所述3D点的对角的大小以及所述3D点的投影误差指示所述3D点足够稳固用于包含时，将所述3D点包含于所述云中，其中所述对角的所述大小以及所述投影误差与多对渐小相应阈值中的一或多个进行测试，以确定所述3D点是否足够稳固，其中一对阈值由角大小阈值及投影误差阈值组成。

22. 根据权利要求21所述的方法，其中所述多对渐小相应阈值中的初始阈值对包含基准角大小阈值以及基准投影误差阈值，且通过将对应的预定渐大缩放因子应用于所述基准角大小阈值以及所述基准投影误差阈值来确定所述多对渐小相应阈值中的其它阈值的阈值。

23. 根据权利要求22所述的方法，其中有五对渐小相应阈值，且所述渐大缩放因子为2、4、8以及16。

24. 根据权利要求21所述的方法，其中N个连续帧的滑动窗口中的两个最近帧用于点对应，以估计所述多个轨迹中的所述2D位置，且每个轨迹包含至多N个2D位置，其中所述多个连续帧中的帧的数目为N，且 $N > 2$ 。

25. 根据权利要求24所述的方法，其中 $N = 6$ 。

26. 根据权利要求21所述的方法，其进一步包括：

将所述云与3D点的先前云合并以产生3D点的最终云，其中，针对所述先前云中的每个3D点，确定是否在所述最终云中包含所述3D点包括：

在所述多个连续帧中的最近帧与所述多个连续帧中的次最近帧之间执行点对应，以尝试定位在对应于第二2D点的所述最近帧中的第一2D点，所述第二2D点在对应于所述3D点的所述次最近帧中；

如果定位所述第一2D点，那么计算所述第一2D点的极线以及所述第二2D点的极线；以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点包含于所述最终云中。

27. 根据权利要求21所述的方法，其进一步包括：

将所述云与3D点的先前云合并以产生3D点的最终云，其中，针对所述先前云中的每个3D点，确定是否在所述最终云中包含所述3D点包括：

在所述多个连续帧中的最近帧与所述多个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所述次最近帧中;

如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点不包含于所述最终云中。

28. 根据权利要求21所述的方法,其中所述点为关注点。

29. 根据权利要求21所述的方法,其中所述计算机视觉系统包含于汽车安全系统中。

30. 一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中基于由单目相机捕获的多个连续帧产生三维3D点的云;以及在所述计算机视觉系统中将3D点的所述云与3D点的先前云合并以产生3D点的最终云,其中,针对所述先前云中的每个3D点,确定是否在所述最终云中包含所述3D点包括:

在所述多个连续帧中的最近帧与所述多个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所述次最近帧中;

如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点不包含于所述最终云中。

31. 根据权利要求30所述的方法,其中产生3D点的所述云包括:

接收所述多个连续帧中的点的所估计二维2D位置的多个轨迹;以及

通过估计对应于所述多个轨迹中的每个轨迹的3D点来计算3D点的所述云,

其中N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计所述轨迹中的所述2D位置,且每个轨迹包含至多N个2D位置,其中所述多个连续帧中的帧的数目为N,且 $N > 2$ 。

32. 根据权利要求31所述的方法,其中 $N = 6$ 。

33. 根据权利要求30所述的方法,其中产生3D点的点云包括:

接收所述多个连续帧中的点的所估计二维2D位置的多个轨迹;以及

针对所述多个轨迹中的每个轨迹,通过以下操作产生三维3D点的云;

基于所述轨迹中的所述2D位置来三角测量三维3D点的所估计位置;以及

当所述3D点的对角的大小以及所述3D点的投影误差指示所述3D点足够稳固用于包含时,将所述3D点包含于所述云中,其中所述对角的所述大小以及所述投影误差与多对渐小相应阈值中的一或多个进行测试,以确定所述3D点是否足够稳固,其中一对阈值由角大小阈值及投影误差阈值组成。

34. 根据权利要求33所述的方法,其中所述多对渐小相应阈值中的初始阈值对包含基准角大小阈值以及基准投影误差阈值,且通过将对应的预定渐大缩放因子应用于所述基准角大小阈值以及所述基准投影误差阈值来确定所述多对渐小相应阈值中的其它阈值对的阈值。

35. 根据权利要求34所述的方法,其中有五对渐小相应阈值,且所述渐大缩放因子为2、

4、8以及16。

36.一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中产生由单目相机捕获的N个连续帧中的点的所估计二维2D位置的多个轨迹;

通过估计对应于所述多个轨迹中的每个轨迹的3D点而在所述计算机视觉系统中计算三维3D点云,其中N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计所述轨迹中的所述2D位置,且每个轨迹包含至多N个2D位置,其中 $N>2$ ;

将所述3D点云与先前的3D点云合并以产生最终3D点云,其中,针对所述先前的3D点云中的每个3D点,确定是否在所述最终3D点云中包含所述3D点包括:

在所述N个连续帧中的最近帧与所述N个连续帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一2D点,所述第二2D点在对应于所述3D点的所述次最近帧中;

如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以及

基于所述第一2D点和所述第二2D点到相应极线的欧几里得距离使所述3D点包含于所述最终3D点云中。

37.一种用于计算机视觉系统中的运动结构处理的方法,所述方法包括:

在所述计算机视觉系统中产生由单目相机捕获的N个连续帧中的点的所估计二维2D位置的多个轨迹;

通过估计对应于所述多个轨迹中的每个轨迹的3D点而在所述计算机视觉系统中计算三维3D点云,其中N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计所述轨迹中的所述2D位置,且每个轨迹包含至多N个2D位置,其中 $N>2$ ;以及

其中计算所述3D点云包括:

针对所述多个轨迹中的每个轨迹,通过以下操作产生所述3D点云;

基于所述轨迹中的所述2D位置来三角测量三维3D点的所估计位置;以及

当所述3D点的对角的大小以及所述3D点的投影误差指示所述3D点足够稳固用于包含时,将所述3D点包含于所述3D点云中,其中所述对角的所述大小以及所述投影误差与多对渐小相应阈值中的一或多个进行测试,以确定所述3D点是否足够稳固,其中一对阈值由角大小阈值及投影误差阈值组成。

38.根据权利要求37所述的方法,其中所述多对渐小相应阈值中的初始阈值对包含基准角大小阈值以及基准投影误差阈值,且通过将对应的预定渐大缩放因子应用于所述基准角大小阈值以及所述基准投影误差阈值来确定所述多对渐小相应阈值中的其它阈值对的阈值。

39.根据权利要求38所述的方法,其中有五对渐小相应阈值,且所述渐大缩放因子为2、4、8以及16。

## 用于计算机视觉系统中的运动结构处理的方法

### 技术领域

[0001] 本发明大体上涉及计算机视觉系统,且更具体地说涉及计算机视觉系统中的运动结构。

### 背景技术

[0002] 新等级的嵌入式安全系统(被称作高级辅助驾驶系统(ADAS))已引入至汽车中以减小人类操作失误。此类系统可提供功能性,诸如后视相机、电子稳定性控制和基于视觉式行人检测系统。这些系统中的许多使用单目相机且依赖于计算机视觉处理以检测一或多个相机的视场中的物件。运动结构(SfM)处理为此类系统中的重要操作以理解来自单目相机捕获的二维图像的三维环境。

### 发明内容

[0003] 所描述的实例包含用于计算机视觉系统中的实时运动结构(SfM)处理的方法和系统。在一个方面中,一种用于计算机视觉系统中的运动结构处理的方法,所述方法包含:接收由单目相机捕获的连续帧中的点的所估计二维(2D)位置的轨迹;以及由(针对)轨迹中的每一个产生三维(3D)点的云;基于轨迹中的2D位置三角测量3D点的所估计位置;以及当3D点的对角的大小以及3D点的投影误差指示所述3D点足够稳固用于包含时,将所述3D点包含于云中。所述对角的大小以及所述投影误差与一或多对渐小相应阈值进行测试,以确定所述3D点是否足够稳固。一对阈值由角大小阈值和投影误差阈值组成。

[0004] 在一个方面中,一种用于计算机视觉系统中的运动结构处理的方法包含:基于由单目相机捕获的连续帧产生三维(3D)点的云;以及将所述云与3D点的先前云合并以产生3D点的最终云。针对所述先前云中的每个3D点,确定是否在最终云中包含3D点包含:在连续帧中的最近帧与所述帧中的次最近帧之间执行点对应,以尝试定位在对应于第二2D点的所述最近帧中的第一二维(2D)点,所述第二2D点在对应于所述3D点的次最近帧中;如果定位所述第一2D点,那么计算所述第一2D点的极线以及所述第二2D点的极线;以及使所述3D点包含或不包含于第一2D点和第二2D点到相应极线的基于最终云的欧几里得(Euclidean)距离。

[0005] 在一个方面中,一种用于计算机视觉系统中的运动结构处理的方法包含:接收由单目相机捕获的N个连续帧中的点的所估计二维(2D)位置的轨迹;以及通过估计对应于所述轨迹中的每一个的3D点来计算三维(3D)点云。N个连续帧的滑动窗口中的两个最近帧用于点对应,以估计轨迹中的2D位置,且每个轨迹包含至多N个2D位置,其中 $N > 2$ 。

### 附图说明

[0006] 图1是配置为用于基于单目相机的汽车安全应用的实例多处理器系统芯片(SOC)的高阶方块图。

[0007] 图2是说明经由在图1的实例SOC上实施的实例实时运动结构(SfM)重构管道的数



据流的图。

[0008] 图3是说明图2的SfM重构管道中的滑动窗口处理的实例。

[0009] 图4是用于在图2的SfM重构管道中产生3D点云的方法的流程图。

[0010] 图5是说明所估计3D点的对角的实例。

[0011] 图6是用于确定是否向前携带点云中的3D点的方法的流程图。

## 具体实施方式

[0012] 出于一致性,在各个图中的类似元件由类似的参考数字表示。

[0013] 如上文所述,运动结构(SfM)处理为在许多基于相机的嵌入式安全系统中执行的计算机视觉处理中的重要操作。许多常规SfM方法涉及来自图像收集的三维(3D)重构,其中所述图像为从各种视点且可能通过不同相机采集。康奈尔大学(Cornell University)研发的微软照片旅游和包(Microsoft Photo Tourism and Bundler)为此SfM方法的两个实例。此类SfM方法涉及360度三维(3D)重构,且通常不适用于实时重构。在汽车安全系统中,实时产生图像(帧),且所使用的任何SfM方法需要实时处理图像。实例实施例提供从单目相机接收图像的计算机视觉系统中的实时SfM处理。本文所描述的实时SfM处理利用针对汽车应用采用的3D环境的限定。

[0014] 图1是配置为用于基于单目相机的ADAS的实例多处理器系统芯片(SOC) 100的高阶方块图。在至少一个实例中,SOC 100为购自德克萨斯仪器公司(Texas Instruments, Inc)的TDA3X SOC的实施例。此描述包含SOC 100的组件的高阶描述。实例组件的更详细描述可发现于M.莫迪(M. Mody)等人,“TI的TDA3X平台的高效前置摄像头ADAS应用(High Performance Front Camera ADAS Applications on TI's TDA3X Platform)”中,2015年12月16日到19日于印度班加罗尔(Bangalore, India)的关于高效计算的2015IEEE第22次国际会议的会议记录(Proceedings of 2015IEEE 22nd International Conference on High Performance Computing),第456-463页,以及2014年10月德克萨斯仪器公司的“用于高级辅助驾驶系统(ADAS)技术摘要的TDA3x SOC处理器(TDA3x SOC Processors for Advanced Driver Assist Systems(ADAS) Technical Brief)”SPRT704A,第1-6页,两者以引用的方式并入本文中。

[0015] SOC 100包含经由高速互连件122耦合的双通用处理器(GPP) 102、双数字信号处理器(DSP) 104和视觉处理器106。SOC 100进一步包含直接存储器存取(DMA)组件108、耦合到显示管理组件114的相机捕获组件110、芯片上随机存取(RAM)存储器116以及经由互连件122全部耦合到处理器的各种输入/输出(I/O)周边装置120。此外,SOC 100包含安全组件118,安全组件118包含安全相关功能性以实现遵从汽车安全要求。此类功能性可包含对于数据的CRC(循环冗余检查)的支持、漂移检测的时脉比较器、错误传信、加窗看门狗定时器和SOC对于损坏和故障的自测试。如本文所述实施实时SfM的软件可存储于存储器116中,且可在SOC 100的一或多个可编程处理器上执行。

[0016] 图2是说明经由在图1的实例SOC 100上实施的实施为计算机视觉系统的一部分的实时SfM重构管道的实施例的数据流的图。在接收每个输入帧之后,SfM重构管道产生3D点云。在高层处,SfM重构管道检测并跟踪跨越连续帧的关注点的二维(2D)位置,将2D位置存储于对应于关注点的轨迹中。因此,轨迹为随时间推移的关注点的一系列2D帧坐标。对轨迹

中的每一个执行三角测量以产生3D点云,所述3D点云随后用于产生最终3D点云。

[0017] 此外,如图3中所示,管道以滑动窗口方式操作。窗口包含N个帧,所述N个帧为先前处理的N-1个帧以及当前帧。如将在本文中更详细地解释,每个输出点云是基于当前窗口中的帧。此外,经执行以估计3D点的三角测量使用针对每个关注点的当前滑动窗口中的N个帧中的至多N个2D位置。可使用N的任何适当值,且N的值可选择为点云质量(质量随N的值增加而增加)与如计算复杂度、性能和存储器大小等考虑之间的取舍。为易于描述,在下文描述中假设N=6。

[0018] 参考图2,每次经由捕获组件110从单目相机124接收帧时,对帧(即,帧(t))执行关注点检测200以检测关注点。帧(t)和所检测的关注点存储于存储器201中。举例来说,存储器201可为图1的芯片上RAM 116。可使用用于关注点检测的任何合适技术来执行关注点(也被称为关键点或特征)的检测200。此外,例如,关注点检测可基于哈里斯(Harris)拐点检测或来自加速片段测试(FAST)检测的特征。哈里斯拐点检测描述于C.哈里斯(C.Harris)和M.斯蒂芬斯(M.Stephens),“合并的拐点和边缘检测器(A Combined Corner and Edge Detector)”中,第四次阿尔维视觉会议的会议记录(Proceedings of Fourth Alvey Vision Conference),英国曼彻斯特(Manchester,UK),第147-151页,1988。FAST描述于E.罗森(E.Rosten)和T.杜伦孟德(T.Drummond),“用于高速拐点检测的机器学习(Machine Learning for High Speed Corner Detection)”中,2006年5月7日到13日于奥地利格拉兹(Graz,Austria)的第9次欧洲计算机视觉会议的会议记录(Proceedings of 9th European Conference on Computer Vision)第1卷,第430-443页。

[0019] 在帧(t)中检测到关注点之后,执行当前帧(t)与前一帧(t-1)之间的点对应202。通常,点对应将前一帧(t-1)中找到的关注点及当前轨迹中的每一个与当前帧(t)匹配。在后一种情况下,当前轨迹中的最新2D点用于点对应。当在帧(t-1)与帧(t-2)之间执行点对应时,轨迹中的最新2D点为帧(t-1)中找到的2D点。此外,当前轨迹存储于存储器201中。如果在当前帧(t)中未找到前一帧(t-1)中检测到的关注点的对应,则删除关注点。此外,如果在当前帧(t)中未找到当前轨迹的对应,则删除轨迹。可使用用于点对应的任何合适技术,例如光流技术。举例来说,一些合适的光流技术描述于J.巴伦(J.Barron)等人,“光流技术的性能(Performance of Optical Flow Techniques)”中,国际计算机视觉杂志(International Journal of Computer Vision),第12卷,第1期,第43-77页,1994。

[0020] 随后执行对应验证204以验证前一帧(t-1)与当前帧(t)之间的点对应。为执行验证,基于当前帧(t)与前一帧(t-1)之间的对应的关注点来计算相机基本矩阵。可使用用于计算基本矩阵的任何合适技术。一个实例技术在2016年5月14日描述于[https://en.wikipedia.org/wiki/Eight-point\\_algorithm](https://en.wikipedia.org/wiki/Eight-point_algorithm)可获得的“八点算法(Eight-point Algorithm)”,Wikipedia。基本矩阵随后用于评估点对应的质量,即,验证点对应。使用基本矩阵来验证点对应为已知的,因此本文中不再进一步解释。

[0021] 如果关注点的点对应无效,那么删除关注点。否则,形成新的轨迹,所述轨迹包含关注点的2D位置以及帧(t)中的对应点的2D位置。此外,如果现有轨迹的点对应有效,那么将对应点的2D位置添加到轨迹。否则,删除所述轨迹。此外,精简帧(t)中检测到的关注点。因此,帧(t)中检测到的关注点具有与当前轨迹中的任何2D位置非常接近或完全相同的坐标,删除检测的关注点。

[0022] 轨迹中的2D位置的数目受滑动窗口中的帧的数目限制,即,任何给定轨迹可具有至多六个点坐标,来自先前五个帧的五个坐标以及来自当前帧的坐标。如果在处理帧(t)时在轨迹中已有六个坐标,那么从轨迹丢弃最旧的坐标且添加来自帧(t)的坐标。

[0023] 随后执行相机位姿计算206。相机位姿由包含旋转的六个自由度构成,即,滚动、倾斜和偏转,以及相机关于现实的3D平移。相机位姿计算206的输出为三度旋转向量以及三度平移向量。估计给出现实中的n个3D点的集合及其帧中的对应2D投影的经校准相机的位姿的问题被称为n点透视成像(PnP)。呈旋转和平移形式的相机的位姿可经发现给出n个3D点的集合及其对应的2D图像投影。可使用任何合适PnP方案技术,例如S.Li等人,“n点透视成像问题的稳固O(n)方案(A Robust O(n) Solution to the Perspective-n-Point Problem)”中描述的技术,关于模式分析和机器智能的IEEE学报(IEEE Transactions on Pattern Analysis and Machine Intelligence),第34卷,第7期,2012年7月,第1444-1450页。

[0024] nP技术需要经由SfM重构管道在先前迭代中产生点云。针对在启动时或在先前迭代中产生点云失败之后在管道中处理的第一对帧,无点云可供使用。在一些实施例中,单一值分解(SVD)技术用以此类情况下的位姿估计。SVD技术的输入为两个帧,两个帧之间的关注点对应,以及基本矩阵。在后续迭代中,点云可供使用,且使用PnP技术来执行位姿估计。可使用任何合适的基于SVD的技术。一个实例技术描述于R.哈拉利克(R.Haralick)等人,“根据对应点数据的位姿估计(Pose Estimation from Corresponding Point Data)”中,IEEE系统学报(IEEE Transactions on Systems),人类以及控制论(Man, and Cybernetics),第19卷,第6期,第1426-1446页,1989年11月/12月。

[0025] 随后执行三角测量计算208以产生对应于当前滑动窗口中的帧的点云。图4为由三角测量计算208执行以产生3D点云的方法的流程图。针对每一轨迹410,执行三角测量400以在轨迹中找到对应于2D位置的3D位置。如上文所述,轨迹记录当前滑动窗口中的两个或两个以上连续帧中的关注点的位置。三角测量400使用轨迹中的2D坐标基于当前所估计的相机位姿和当前基本矩阵找到对应的3D位置。由于滑动窗口大小为六个帧,因此关注点的轨迹可具有至多六个2D坐标。

[0026] 可使用任何合适的三角测量技术。举例来说,一些合适的技术描述于R.哈脱莱(R.Hartley)和P.斯特姆(P.Sturm),“三角测量(Triangulation)”中,计算机视觉和图像理解(Computer Vision and Image Understanding),第68卷,第2期,第146-157页,1997年11月(本文中为“哈脱莱(Hartley)”),其通过引用的方式并入本文中。在一些实施例中,使用哈脱莱(Hartley)中描述的迭代最小平方技术。哈脱莱(Hartley)使用两个帧描述三角测量,但哈脱莱(Hartley)中的技术可延伸为超过两个帧。

[0027] 在估计对应于轨迹的3D点之后,将3D点重投影402到用于估计3D位置的两个2D点的帧中,且确定402每个2D点的重投影误差。通常,重投影误差为对应于重投影2D点与对应的所测量2D点之间的欧几里得距离的几何误差。重投影误差定量3D点的估计重新产生所测量2D点的真实投影的接近程度。理想地,重投影误差为零,且非零值指示三角测量的误差。因此,重投影误差可为3D点的稳固性的指示。

[0028] 此外,确定3D点的对角404。图5是说明所估计3D点的对角的实例。在此实例中,关注点的轨迹包含对应于三个连续帧中的每一个中的位置的三个2D点。对角由3D点处的左边

光线和右边光线的交叉点形成。左边光线通过光学中心和帧N中的所测量2D点位置,且右边光线通过光学中心和帧N+2中的所测量2D点位置。由于更小的角可负面影响三角测量的质量,对角的大小可为3D点的稳固性的指示。

[0029] 再次参考图4,确定406是否精简3D点,即,关于是否将3D点包含于点云中。此决策是基于经三角测量以估计3D点的2D点的重投影误差,以及基于所估计3D点的对角的大小。通常,将角大小和重投影误差的总和与对应的凭经验确定的阈值进行比较,且使用比较结果来决定3D点是否足够稳固以包含于点云中。

[0030] 在许多视觉应用中,单一对角阈值和单一投影误差阈值足够用于精简。然而,使用单一阈值在汽车应用中存在问题,尤其在使用前置摄像头的应用中。当车辆以直线向前移动时,对应于车辆前方物件的点的对角可变得极小。当对应于关注物件的点接近光学中心时,此效果更明显。稳固地检测中心的物件对于采取行动避免碰撞此类物件极为重要。测试表明此类汽车应用,将角阈值设定为0.25度精简光学中心附近的太多3D点,这降低了3D点的密集度,而将角阈值设定为0.125度不提供充分精简,因此保留太多不良重构的3D点。为了得到光学中心周围的3D点的良好密集度,同时还精简不良重构的3D点,单一角阈值并不存在。

[0031] 在一些实施例中,投影误差和对角值与所选择的多对渐小相应阈值进行比较,而非使用对角阈值的单一值和投影误差阈值的单一值,以使得当对角的大小较大时,在精简决策中容许更大投影误差,且当对角的大小较小时,在精简决策中容许更小投影误差。使用渐小阈值确保否则的话不精简具有更小对角和可接收投影误差的稳固3D点。

[0032] 表1的实例伪代码说明此精简方法。在此实例伪代码中,subtendedAngle为对角的大小,projectionError为两个2D点的重投影误差的总和,angleTh为基准角大小阈值,projectionErrTh为基准投影误差阈值,且 $k_x, x=1, 2, 3, 4$ 为用于减小angleTh和projectionErrTh的缩放因子。如果对角大小或投影误差不能进行相应阈值比较测试,那么使基准阈值减小下一较大缩放因子,且再次尝试阈值比较测试。如果所有缩放因子均尝试不成功,那么精简3D点。可使用用于基准角大小阈值、基准投影误差阈值和缩放因子的任何合适值。在一些实施例中,凭经验确定这些值。此外,使用更多或更少缩放因子的实施例为可能的。

[0033] 表1

```

k1= 2; k2= 4; k3 = 8; k4 =16
prune = true;
if((subtendedAngle>=angleTh)&&(projectionError<=projectionErrTh)) {
    prune = false;
} else if ((subtendedAngle >= angleTh/k1 && (projectionError <= projectionErrTh/k1)) {
    prune = false;
[0034] } else if ((subtendedAngle >= angleTh/k2) && (projectionError <= projectionErrTh/k2)) {
    prune = false;
} else if ((subtendedAngle >= angleTh/k3) && (projectionError <= projectionErrTh/k3)) {
    prune = false;
} else if ((subtendedAngle >= angleTh/k4) && (projectionError <= projectionErrTh/k4)) {
    prune = false;
}

```

[0035] 再次参考图2,在三角测量计算208之后,执行最终3D点云产生210。通过将通过三

角测量计算208输出的3D点云与针对先前滑动窗口产生的3D点云合并来产生最终3D点云。两个点云的简单合并存在问题,这是因为当前点云中的新的点可能对应于新的物件,所述物件阻隔对应于先前点云中的较旧点的物件。不应将此类较旧的点向前携带至最终的点云,这是因为这些点不代表新的物件。

[0036] 在一些实施例中,在合并两个点云之前,使用基本矩阵关系测试来测试先前3D点云中的3D点以决定是否应将3D点向前携带至最终3D点云中。图6是用于决定是否向前携带3D点的实例方法的流程图。初始地,根据对应于3D点的轨迹来确定600对应于先前3D点云中的3D点的前一帧(t-1)中的2D点。随后使用点对应在当前帧(t)中找到602对应于前一帧(t-1)中的2D点的2D点。可使用用于点对应的任何合适技术,例如上文所述的技术。如果点对应在当前帧(t)中未找到604有效的对应点,那么从先前的3D点去除3D点612,并且继续614下一3D点(如果存在)的处理。

[0037] 如果找到604有效的对应点,那么应用基本矩阵关系测试606-610,以确定是否向前携带3D点。更具体地说,使用基本矩阵来计算606对应于2D点的极线。因此,如果x为前一帧(t-1)中的2D点,x'为当前帧(t)中的对应2D点,且F为基本矩阵,那么 $l' = Fx$ 为对应于x的极线,且 $l = FTx'$ 为对应于x'的极线。随后计算每个2D点到对应的极线的欧几里得距离608。随后将欧几里得距离的总和与阈值进行比较604以确定是否向前携带3D点。如果总和足够小,即小于或等于阈值,那么向前携带3D点。否则,从先前的点云去除3D点612。在任一情况下,继续614下一3D点(如果存在)的处理。任何合适的值可用于所述阈值。可凭经验确定所述值。

[0038] 其它实施例

[0039] 本文已描述使用PnP方案技术来计算相机位姿估计的实例实施例。在至少一个实例实施例中,惯性测量单元(IMU)用于提高使用PnP方案技术发现的位姿估计,或IMU替换相机位姿估计。

[0040] 在另一实例中,本文已描述执行关注点检测的实施例。在至少一个实例实施例中,未执行关注点检测,且针对帧中的所有像素(点)执行点对应。

[0041] 在至少一个实例实施例中,将用于光束平差的合适技术添加到SfM重构管道以用于相机位姿估计和三角测量的共同优化。

[0042] 在另一实例中,在本文中已描述在SfM重构管道中执行的处理被实施为在多处理器SOC中的处理器上执行的软件指令的实施例。SfM重构管道的功能性可被实施为软件、固件和硬件的任何合适的组合。举例来说,功能性中的一些可实施在一或多个硬件加速器、专用集成电路(ASIC)和/或现场可编程门阵列(FPGA)中。

[0043] 在另一实例中,本文中已描述实施例,其中在先前点云中的所有3D点上执行结转处理以去除任何3D点,所述3D点在合并先前点云和当前点云之前不应被结转。在又另一实例中,当两个点云合并时,本文所述的结转决策标准应用于先前点云中的点。

[0044] 在另一实例中,已参考汽车安全系统在本文中描述了实施例。实例实施例针对具有类似环境限定的其它计算机视觉应用为可能的,例如工业应用,机器人技术和消费应用(例如,真空清洁器和无人机)。

[0045] 尽管可在本文中以依序方式呈现且描述方法步骤,但是在图所展示且在本文中描述的步骤中的一或多者可并行地予以执行、可组合、及/或可按不同于在图所展示且

在本文中所描述的次序的次序予以执行。因此,实施例不限于图所展示及/或本文中所描述的步骤的特定次序。

[0046] 在本说明书中,术语“耦合”及其派生词意指间接、直接、光学及/或无线电气连接。举例来说,如果第一个装置耦合到第二个装置,那么那个连接可以通过直接电连接、通过经由其它装置和连接的间接电连接、通过光学电连接和/或通过无线电连接。

[0047] 在所描述的实施例中可能进行修改,且其它实施例在权利要求的范围内为可能的。

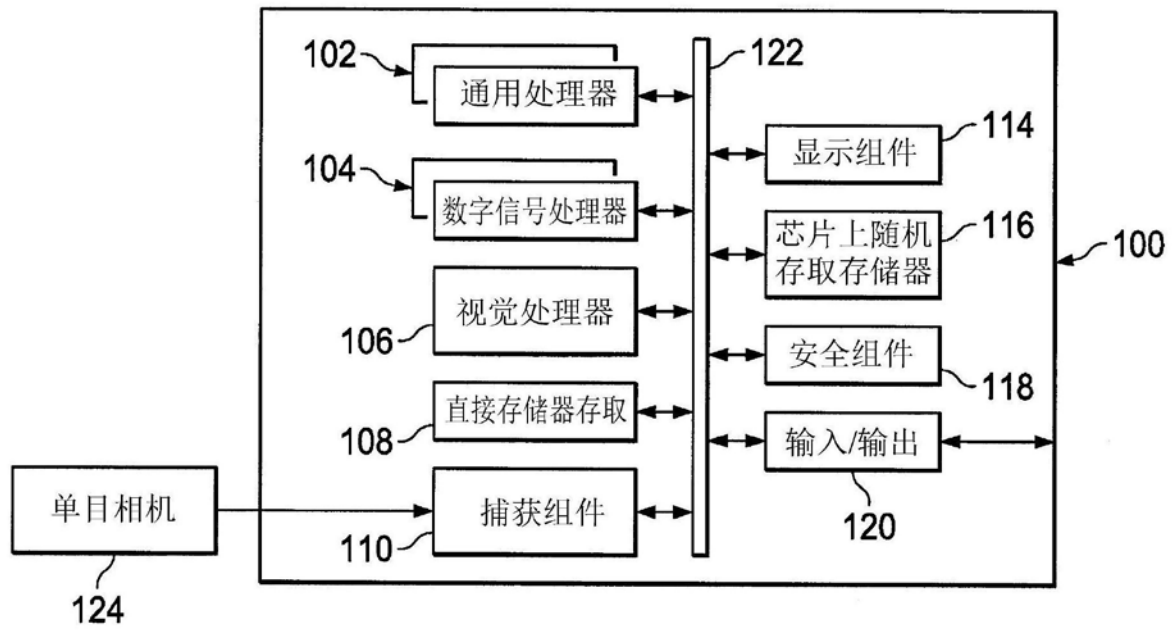


图1

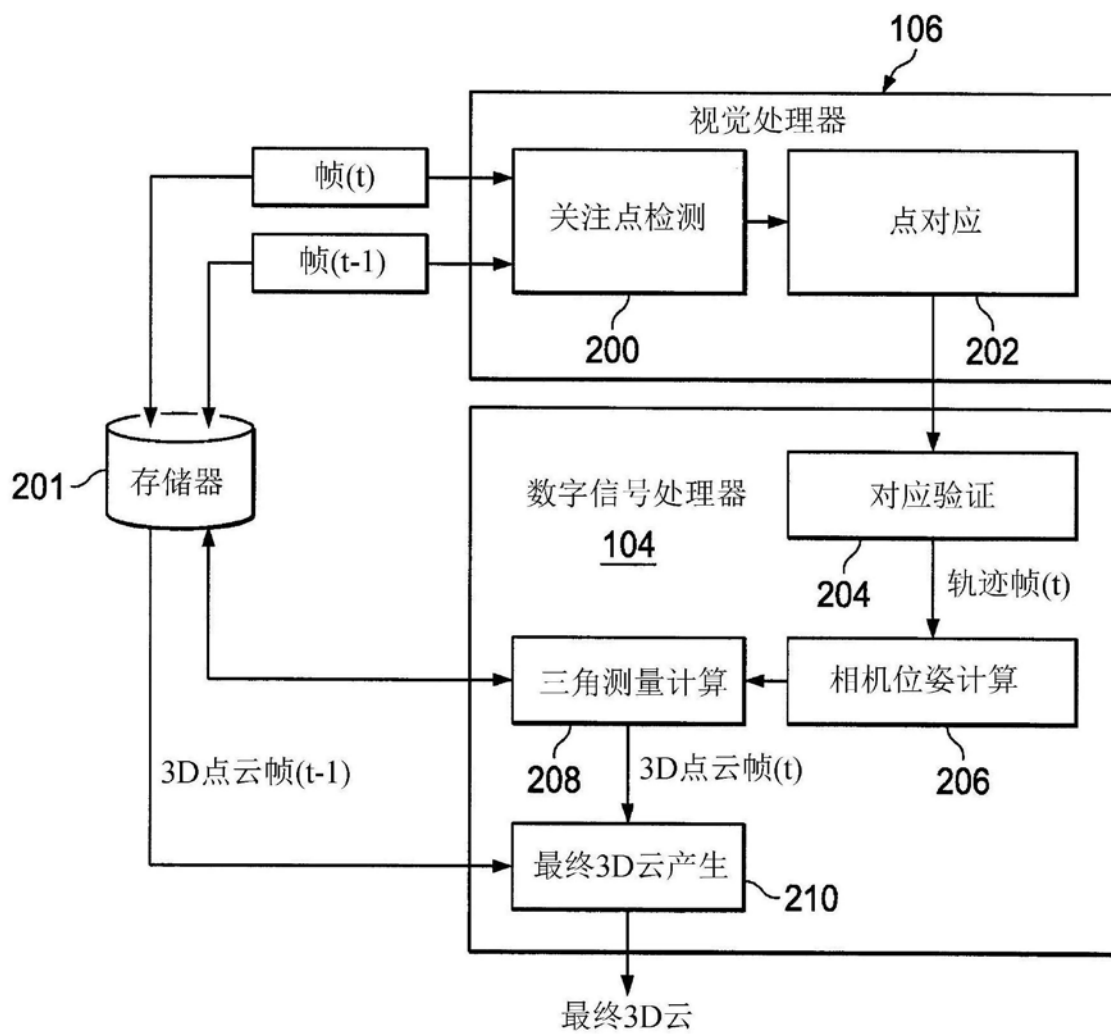


图2



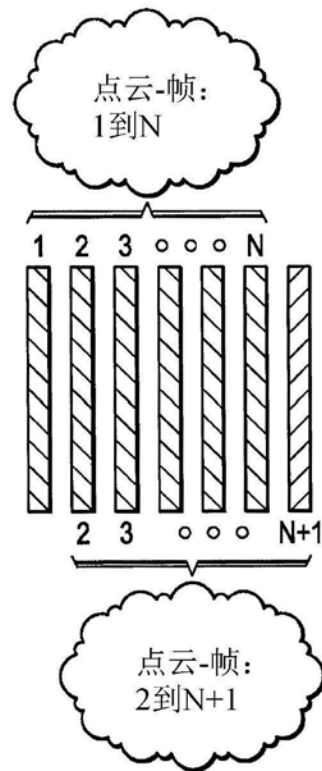


图3

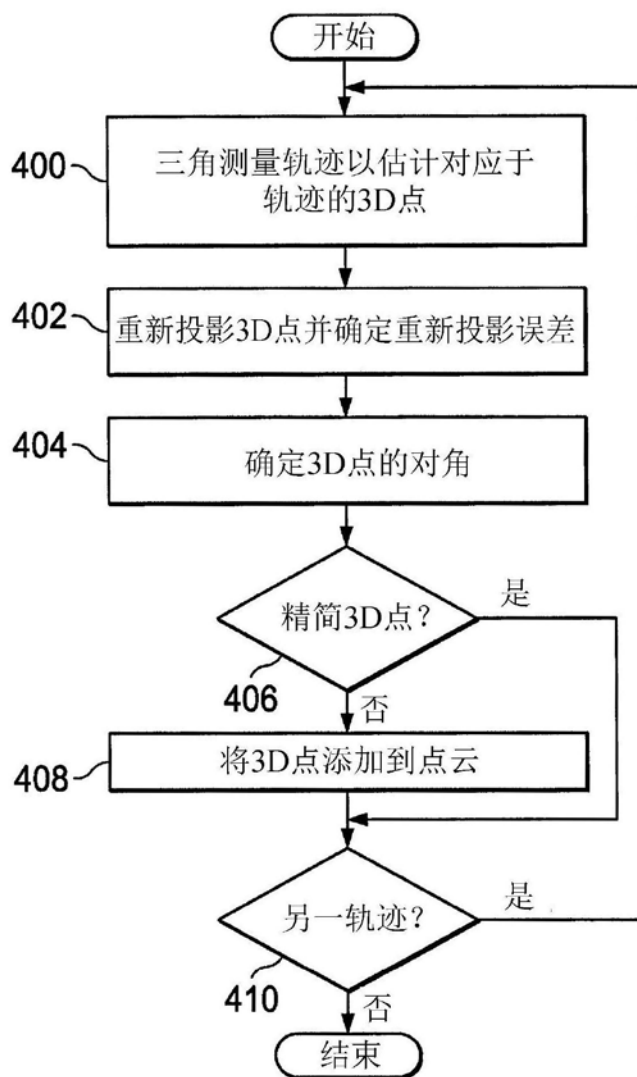


图4

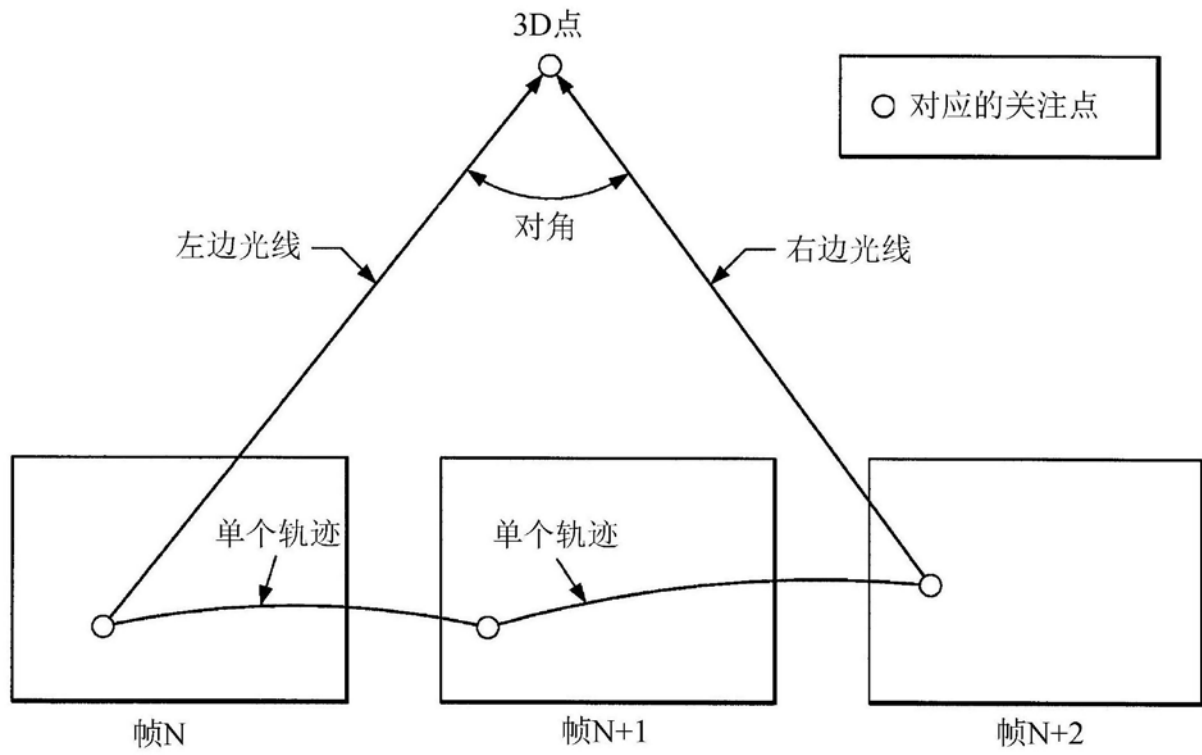


图5

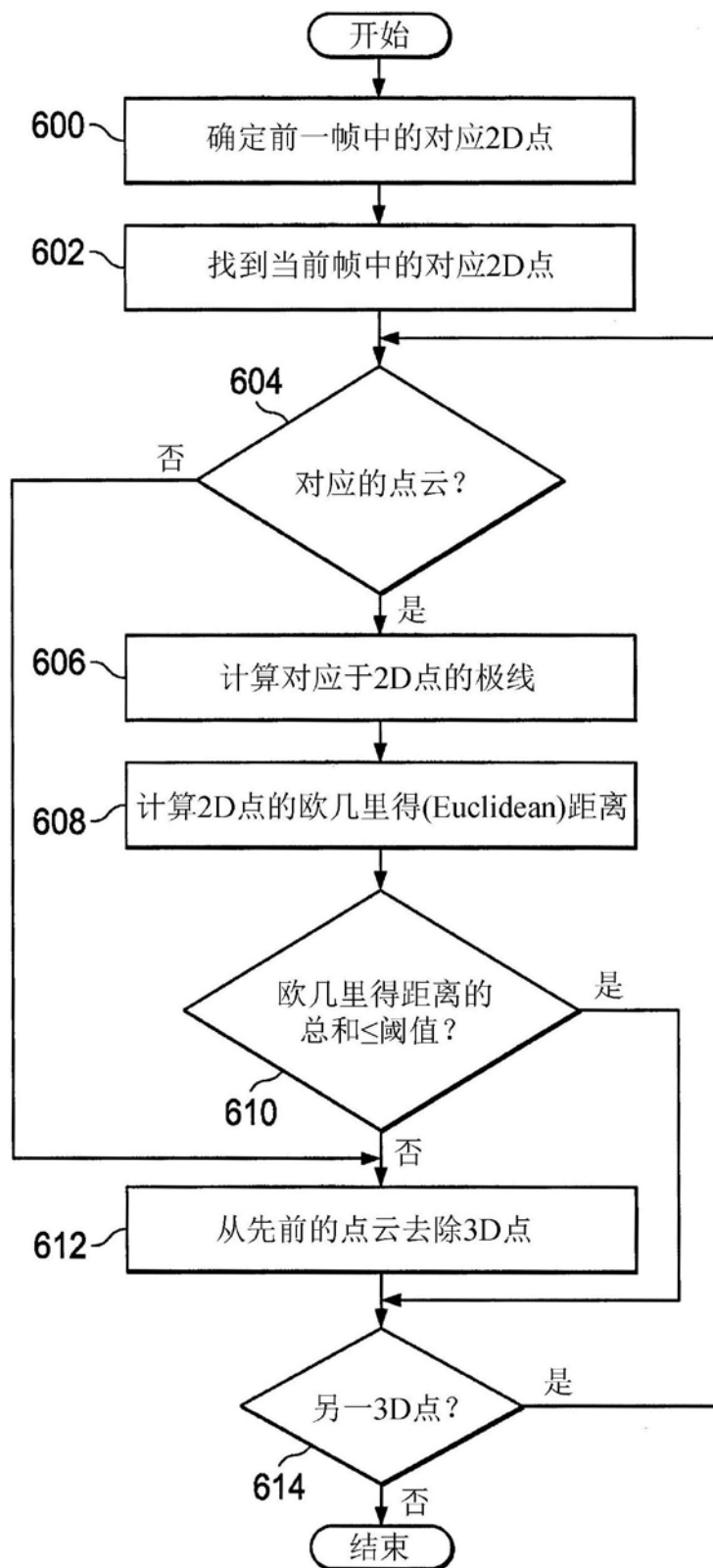


图6