引文格式：何梁, 黄劲松. 一种基于去运动模糊的视觉SLAM算法[J]. 地理空间信息, 2023, 21(3): 31-35. doi:10. 3969/j.issn. 1672-4623. 2023.03.007

一种基于去运动模糊的视觉SLAM算法

何　梁[1]，黄劲松[1]
（1. 武汉大学 测绘学院，湖北 武汉 430079）

摘　要：针对视觉SLAM中因相机高速运动带来的图像运动模糊导致定位精度下降的问题，采用一种基于生成对抗网络的去运动模糊方法进行图像的运动模糊校正，将其应用于视觉SLAM算法中，并提出一种基于滑动窗口的关键帧筛选与维护方法。在数据集中进行测试，结果表明该关键帧维护方法可有效减小计算量，提高计算效率。结合去运动模糊的视觉SLAM算法可提高特征点提取与匹配的效果，提升在相机剧烈运动状况下SLAM的精度。

关键词：去运动模糊；视觉SLAM；生成对抗网络；关键帧
中图分类号：P232　　**文献标志码：**B　　**文章编号：**1672-4623（2023）03-0031-05

Visual SLAM Algorithm Based on Demotion Fuzzy

HE Liang[1], HUANG Jinsong[1]
(1. School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China)

Abstract: Aiming at the problem of image motion fuzzy caused by high speed camera motion in visual SLAM decreasing the localization accuracy, we used a demotion fuzzy method based on generative adversarial network to correct the motion fuzzy of image, applied it to visual SLAM algorithm, and proposed a key frame screening and maintenance method based on sliding window. We tested it in a data set. The results show that the key frame maintenance method can effectively reduce the computation amount and improve the computation efficiency. The visual SLAM algorithm combined with demotion fuzzy can improve the effect of feature point extraction and matching, and improve the accuracy of SLAM algorithm under the condition of camera intense motion.

Key words: demotion fuzzy, visual SLAM, generative adversarial network, key frame

SLAM[1-2]技术被广泛应用于自动驾驶、AR/VR以及移动测量[3]等领域。随着计算机视觉的迅猛发展，基于视觉的SLAM方法成为研究热点，主流算法有ORB-SLAM等。视觉SLAM所需的传感器为相机，然而在相机高速运动时经常出现运动模糊的场景。运动模糊是指景物图像中像素发生了移动和叠加。若在曝光期间相机和物体发生了速度较快的相对移动，就会出现比较明显的运动模糊。运动模糊是降低视觉SLAM性能的重要因素，因此有必要对其进行研究。

在去运动模糊领域，LI Y[4]等结合维纳滤波，利用一种新的模糊核估计模型实现对旋转运动模糊图像的去模糊，达到了良好效果；Choudhury S[5]等提出了一种加性高斯噪声辅助的去运动模糊方法，在算法中利用正则化函数取代固定参数使得收敛速度更快；XU L[6]提出了一种基于L0范数的最大后验概率去运动模糊方法；SUN J[7]等先将深度学习方法应用于图像去模糊领域，再采用卷积神经网络进行模糊核估计，而不是直接端到端地去模糊，最后仍采用非盲去模糊反卷积对图像进行还原；Nah S[8]等提出的卷积神经网络可通过端到端的方式实现去模糊，在视觉效果和定量分析上都取得了令人满意的表现；Kupyn O[9]等提出了基于条件生成对抗网络的图像复原算法DeblurGAN，可对单张图片直接进行去模糊，达到了较佳的效果。在去运动模糊与SLAM相结合方面，Lee H S[10]等通过来自SLAM系统的信息来估计运动模糊，再利用估计值去模糊；JIN H[11]等提出了不同思路，不是去模糊，而是对图像块进行模糊处理，使其看起来相似，这可能有助于特征匹配，但在存在运动模糊的情况下，检测到的特征数量通常会大大减少，特别是当显著特征的数量有限时，将对系统精度造成影响；Pretto A[12]等针对许多分割的图像区域估算了点扩散函数（PSF），用于最小化运动模糊在基于SIFT提取兴趣点和构建图像描述符时的不良影响，该方法无需显式的去模糊，但计算时间不足以进行实时操作，且基于反卷积的去模糊

收稿日期：2021-05-28；修回日期：2022-09-15。
第一作者简介：何梁（1996—），硕士研究生，研究方向为导航与定位技术，E-mail：441384003@qq.com。

可能会使图像变差，不是解决视觉SLAM中运动模糊的适当方案，因为恢复的图像质量很大程度上取决于估计的PSF准确性；Klein G[13]等试图使用Edgelets解决视觉SLAM中的模糊效果，Edgelets是指局部直线段，即使在严重模糊的图像中，小边缘也可能保持完整，据此提出了一种使用Edgelets的跟踪方法，并使视觉SLAM系统对运动模糊具有鲁棒性。

受上述研究启发，由于神经网络能更好地刻画图像特征，将其应用于图像去模糊比传统方法具有更好的效果，因此本文将基于生成对抗网络的去模糊方法与视觉SLAM相结合；由于增加了去运动模糊模块，为提高计算效率，提出了基于特征点匹配数量的运动模糊判断方法，结合局部地图的关键帧选取方法和基于滑动窗口的关键帧维护方法，提升了算法的计算效率和定位精度。

1 基于去运动模糊的视觉SLAM算法

基于去运动模糊的RGB-D相机视觉里程计算流程如图1所示：①对输入的当前帧进行畸变校正；②提取当前帧的关键点并计算描述子；③判断系统是否初始化，未初始化则根据该帧和RGB-D相机提供的深度图计算特征点的三维位置，完成初始化，然后返回步骤②；④判断运动模糊帧和进行去运动模糊，算法分为两个部分，一部分是在特征点匹配数量足够的情况下进行正常的前端跟踪，另一部分是对特征点匹配数量不足的帧进行去运动模糊，再为全局优化提供信息；⑤若经过去模糊后的图像位姿计算成功，则采取结合局部地图的判断策略判断其是否为关键帧，如果是则送入后端优化模块，若位姿仍计算失败则丢弃该帧，由于只在后端优化模块中进行处理，因此无需较高的更新频率。

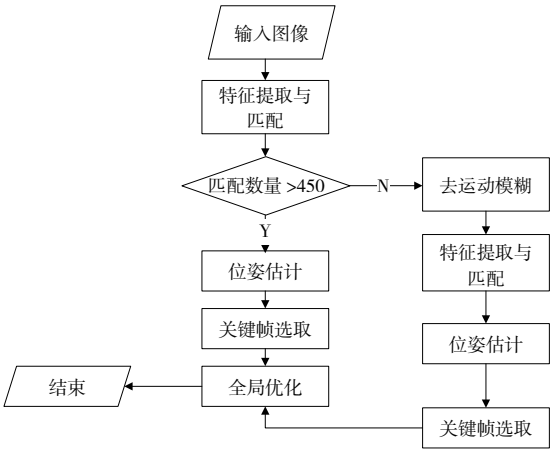


图1 基于去运动模糊的视觉SLAM系统流程图

1.1 基于生成对抗网络的去运动模糊

生成对抗网络的提出使神经网络在图像处理领域得到了更广泛的应用。其应用在去运动模糊上具有更佳的效果，本文基于DeblurGAN的开源算法进行去运动模糊。生成对抗网络包括生成模型G和判别模型D两个由卷积神经网络构成的模型，这两个网络就是博弈双方。生成器通过输入随机噪声输出虚假图像，判别器输入的是真实图像和生成器生成的虚假图像，然后判别输入图像是真实图像还是虚假图像。当生成器生成的图像与原图的分布完全重合，判别器也无法区分生成样本与真实样本，即对真实样本和虚假样本的分类准确率达到50%，此时网络训练完成。生成对抗网络的去模糊过程如图2所示。

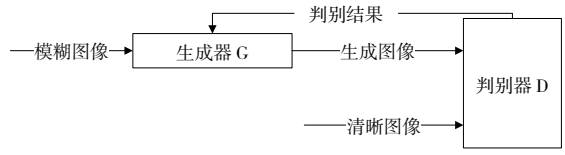


图2 生成对抗网络的去模糊过程示意图

优化函数为：

$$V(D,G)=E_{x\sim p_{\text{data}}(x)}\left[\log D(x)\right]+E_{z\sim p_z(z)}\left[\log\left(1-D\left(G(z)\right)\right)\right] \tag{1}$$

式中，x为训练的真实图片数据；$p_{\text{data}}(x)$为图片数据的分布；$p_z(z)$为噪声z的分布；G为生成器网络，z为随机噪声，通过该噪声生成的图片为$G(z)$；D为判别器网络，判断输入图片是否与真实图片一致；$D(x)$为x是真实图片的概率，1表示真实图片，0表示虚假图片。

训练完毕后，利用训练好的生成器网络可直接对单张图像进行去模糊。

1.2 关键帧选取与维护方法

关键帧是SLAM中具有代表性的帧，能代表附近的帧，从而减少优化帧数。关键帧是连接前端与后端的纽带，合理选取关键帧能平衡精度和计算量。ORB-SLAM2中对于关键帧的判断标准较为宽松，仅将图像的移动距离作为判断标准，即比较当前帧与参考帧的位姿，若超过一定阈值则视为关键帧，因此插入局部地图中的关键帧较多，造成冗余，且插入的关键帧中也容易存在位姿计算不准确的情况。由于本文算法结合了去运动模糊，增加了计算负担，因此需要提高计算效率。针对该问题，本文采取的关键帧选取策略为：每个关键帧主要存储前端普通帧和去运动模糊帧解算的位姿、提取的特征点和匹配的地图点。

由于本文对部分模糊图像进行了去模糊处理，因

此先将去模糊后的图像均视为潜在关键帧。对图像进行去模糊虽然增加了特征点提取的数量，但匹配时无法避免地会增加误匹配情况，本文采用随机采样一致性（RANSAC）算法处理误匹配，使匹配效果更鲁棒。RANSAC算法是一种根据带有外点的实际数据来估计数学模型参数的非确定性算法。根据RANSAC算法去除误匹配的点后，再计算位姿，若外点比例超过一定阈值则不设为关键帧，最后基于以下两点来判断当前帧是否为关键帧：①将当前去模糊帧的位姿与局部地图中姿态最接近或距离最接近的关键帧进行比较，而不是与参考帧的位姿比较，若大于一定阈值则将该帧视为关键帧；②计算当前去模糊帧匹配的地图点数量与当前帧提取的特征点数量之比，若小于一定阈值，则视为关键帧。标准①顾及了局部地图，能利用更多的过往帧信息，标准②则是判断场景是否发生了明显变化，这时也需要更新局部地图。根据上述两点即可筛选出质量较高的关键帧。

关键帧的维护是指在局部地图中适当删除某些帧，防止数据冗余，保证局部地图的规模。在ORB-SLAM2中选取具有共视关系的关键帧加入局部地图，且没有设计关键帧剔除策略，因此会导致关键帧冗余，增加计算量。本文针对这一点做出改进，在局部地图模块中采用滑动窗口的方法，滑动窗口中只保留固定大小的数据，新的关键帧产生后，进行数据关联，将与最新两个关键帧视场变化最大的关键帧删除。具体实现方法为将滑动窗口中每一个关键帧对应的地图点与当前最新两帧的地图点进行匹配，选择出匹配数量最少的关键帧进行剔除，以此进行关键帧的维护。

2　算法测试与分析

2.1　去运动模糊实验与分析

2.1.1　训练过程与设备介绍

训练网络的数据集采用GOPRO数据集。GOPRO数据集是由Nah S[8]等提出的，在图像质量分析中使用最广泛的公开数据集之一。本文采用PyTorch训练模型，将其中的图像随机裁剪成256×256的图像块输入网络训练，模型训练时一次训练所选取的样本数为1；最初生成器和判别器设置的学习速率为10^{-4}，经过150次迭代后，在新一轮150次迭代中将该比率线性衰减；运行电脑的CPU为Intel 酷睿i7 8750H，显卡为NVIDIA GeForce GTX 1060。

2.1.2　去运动模糊效果分析

本文利用未经训练的SLAM数据集TUM数据集进行实验测试。TUM数据集是德国慕尼黑工业大学计算机视觉组2012年提出的一个RGB-D数据集，主要是针对纹理丰富的办公室场景，从中选取一些在相机高速运动场景拍摄的模糊影像进行去模糊测试。

本文选取的去模糊前后的对比示例图如图3、4所示。除直观感受外，本文采用基于Laplacian梯度函数的图像清晰度进行评估。Laplacian梯度函数是一种无需参考图像的图像模糊评价指标，采用Laplacian算子分别提取水平和垂直方向的梯度值。该算子的定义为：

$$\boldsymbol{L}=\frac{1}{6}\begin{bmatrix}1 & 4 & 1\\ 4 & -20 & 4\\ 1 & 4 & 1\end{bmatrix} \tag{2}$$

本文将基于Laplacian梯度函数的图像清晰度定义为：

$$D(f)=\sum_y\sum_x|G(x,y)| \quad G(x,y)>T \tag{3}$$

式中，$G(x,y)$为像素点(x,y)处Laplacian算子的卷积。



图3　效果图1

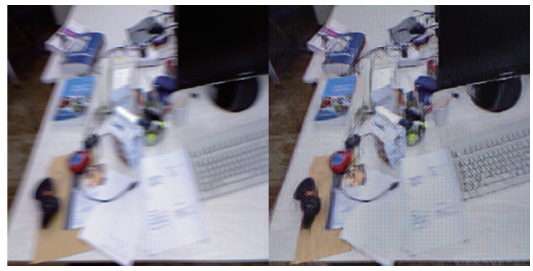


图4　效果图2

Laplacian梯度函数值越大表示图像越清晰，该评价标准与主观感受也较一致。去模糊前后清晰度对比结果如表1所示，可以看出，本文将去运动模糊方法应用于SLAM数据集取得了较好效果。

表1　TUM数据集图像去模糊前后清晰度对比

实验序列	去模糊前清晰度	去模糊后清晰度
第一组	5.278	15.349
第二组	5.261	14.329

2.2　基于去运动模糊的视觉SLAM运行实验与分析

2.2.1　去运动模糊对特征匹配的影响分析

将本文算法跟踪阶段前后帧特征匹配的数量与未进行去运动模糊的SLAM算法进行对比，采用TUM数

据集中的rgbd_dataset_freiburg1_xyz数据包。该段数据采用ORB特征匹配方法进行匹配，设置匹配阈值为500，而模糊图像前后帧匹配数量明显少于500个，因此将匹配数量小于450个的前后帧图像视为模糊图像，均采用神经网络进行去运动模糊。去运动模糊前后图像与上一帧特征点的匹配数量对比如图5所示。为了更直观地展示，本文绘制了去模糊前后匹配数量的增减图（图6），可以看出，大部分图像去模糊后与下一帧的匹配数量均有所增加；该段数据共791帧，其中运动模糊明显的图像有324帧，图像去模糊后提升了与下一帧的匹配效果。

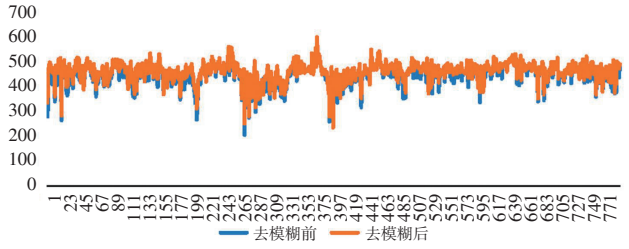


图5 去模糊前后匹配数量对比图

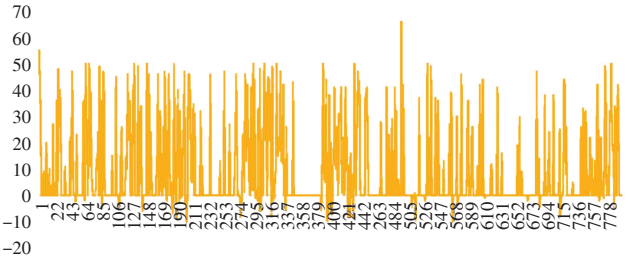


图6 去模糊前后匹配数量增减图

由上述实验可知，通过判断特征数量的方法来判断运动模糊具有有效性，且具有实用价值；每帧去运动模糊的平均时间为315 ms，为了提高计算效率，减少每帧的处理时间十分重要，由于特征点法SLAM中特征提取与匹配是必要的一步，因此通过特征数量来判断运动模糊的方法相较于其他需要计算某个指标再去判断运动模糊的方法，节省了步骤和时间。

2.2.2 算法效率对比

本文将使用本文关键帧选取与维护方法后的计算效率与ORB-SLAM2算法的计算效率进行对比，先分别统计两种算法运行时每个时刻关键帧的数量，再绘制随时间变化的曲线，如图7所示，可以看出，由于本文采用了基于滑动窗口的关键帧维护方法，滑动窗口大小为10帧，因此局部地图中插入的关键帧维持在固定规模，当新的关键帧插入后才会剔除旧的关键帧；而ORB-SLAM2算法选取的是具有共视关系的关键帧插入局部地图，导致关键帧的数量较多。

本文分别计算了两种算法每帧运行耗时，ORB-SLAM2算法为47.4 ms/帧，本文算法为42.7 ms/帧，效率提升了9.9%，可以看出，本文算法由于采取了合理的关键帧维护方法，有效控制了局部地图的规模，虽然包含了去运动模糊模块，但效率仍比ORB-SLAM2算法有所提升。

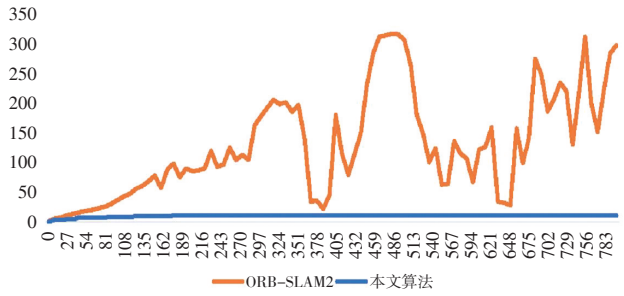


图7 本文算法与ORB-SLAM2算法局部地图中关键帧数量对比

2.2.3 算法精度分析

为评估本文SLAM算法的性能，本文采用SLAM公开数据集的TUM数据集对算法进行测试。数据集中也包含真值，可用于算法效果的比较与测试。本次实验采用其中的rgbd_dataset_freiburg1_xyz数据包，该段数据相机运动速度较快，包含部分运动场景。

本文采用基于去运动模糊的视觉SLAM系统在TUM数据集中进行测试，与未进行去运动模糊的视觉里程计运行结果以及真值进行比较；采用TUM数据集中的rgbd_dataset_freiburg1_xyz数据包进行实验测试，绘制了轨迹对比图（图8）以及位置和位姿解算对比图（图9、10），其中blur表示无去运动模糊模块的视觉里程计的解算结果，result表示基于去运动模糊的视觉SLAM系统的解算结果，truth为真值。

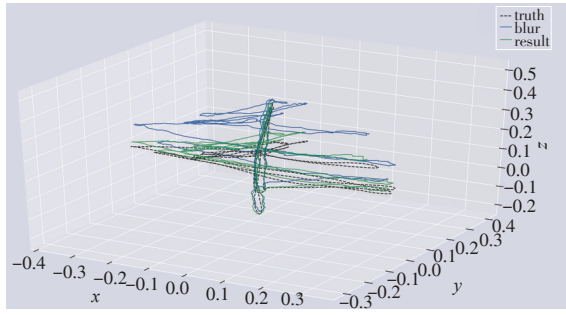


图8 全局轨迹对比图/m

由图8可知，基于去运动模糊的视觉SLAM算法运行所得的轨迹与真值更接近，尤其是在Z方向上；由图9可知，基于去运动模糊的视觉SLAM算法在X和Z方向上的位置解算精度都均有所提升，且在Z方向上的提升效果明显；由图10可知，各个方向上的姿态解算结果均有所提升，由于基于去运动模糊的视觉

SLAM算法优化了前端特征提取匹配效果，并在后端模块插入了更多有效关键帧，使得SLAM系统的精度整体得到了提升。

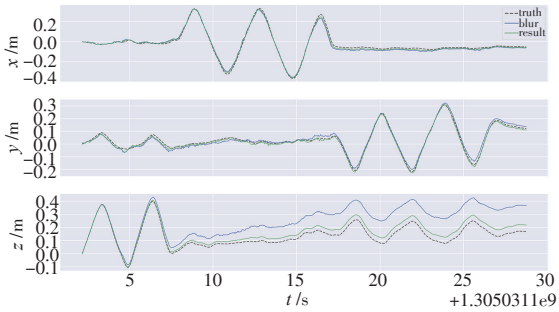


图9　位置解算对比图

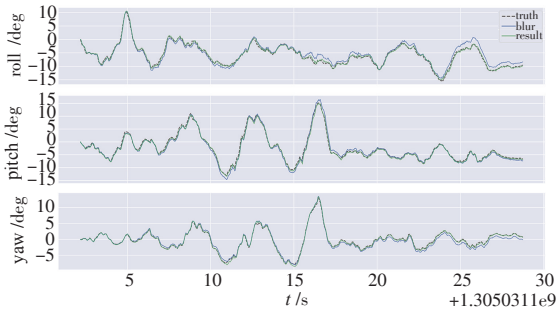


图10　姿态解算对比图

本文分别对去运动模糊前后的绝对位置误差进行了计算分析，得到平均值（Mean）、最大误差（max）、标准差（STD）和均方根误差（RMSE）如表2所示，可以看出，基于去运动模糊的视觉SLAM算法的各项误差均小于未进行去运动模糊的视觉里程计，其中RMSE减小了71.6%，精度提升明显，说明本文方法能有效减弱相机快速移动和旋转的影响，提升定位精度。

表2　误差分析表/m

图像类型	Mean	max	STD	RMSE
去模糊前	0.094	0.176	0.056	0.109
去模糊后	0.018	0.034	0.007	0.031

3　结　语

针对视觉SLAM中高速相机产生的运动模糊带来的定位精度下降的问题，本文提出了一个改进算法；采用基于生成对抗网络的去运动模糊方法、合理的关键帧选取方法和基于滑动窗口的关键帧维护方法，在TUM数据集中进行实验，对其中的运动模糊图像进行运动模糊校正，再将去运动模糊与视觉SLAM相结合，分析比较了去运动模糊前后特征点提取与匹配的效果以及SLAM系统的计算效率和精度。结果表明，基于去运动模糊的视觉SLAM算法在相机剧烈运动的状态下仍具有较高的精度，且保证了合格的计算效率。

参考文献

[1] Aulinas J, Petillot Y, Salvi J, et al. The SLAM Problem: a Survey[C]//Proceedings of 11th International Conference of the Catalan Association for Artificial Intellgengce, 2008: 363-371

[2] 高翔, 张涛. 视觉SLAM十四讲:从理论到实践[M]. 北京:电子工业出版社, 2017

[3] 李德仁. 移动测量技术及其应用[J]. 地理空间信息, 2006, 4(4):1-5

[4] LI Y, JIA F, JIANG X, et al. A Proportional Algorithm for Rotary Motion Blurred Images Restoration[C]//Proceedings of 2nd International Conference on Information Systems and Computer Aided Education, Dalian, 2019:288-292

[5] Choudhury S, Sa P K, Padhy R P, et al. A Denoising Inspired Deblurring Framework for Regularized Image Restoration[C]//Proceedings of IEEE Annual India Conference, 2016:1-6

[6] XU L. Unnatural l0 Sparse Representation for Natural Image Deblurring[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2013:1 107-1 114

[7] SUN J, CAO W F, XU Z B, et al. Learning a Convolutional Neural Network for Non-uniform Motion Blur Removal[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos, 2015:769-777

[8] Nah S, Kim T H, Lee K M, et al. Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring[J]. CVPR, 2017(35):257-265

[9] Kupyn O, Mykhailych B. DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2018:8 183-8 192

[10] Lee H S, Kwon J, Lee K M. Simultaneous Localization Mapping and Deblurring[C]//Proceedings of the IEEE International Conference on Computer Vision, 2011:1 203-1 210

[11] JIN H, Favaro P, Cipolla R. Visual Tracking in the Presence of Motion Blur[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2005:18-25

[12] Pretto A, Menegatti E, Bennewitz M. A Visual Odometry Framework Robust to Motion Blur[J]. Robotics and Automation, 2009(14):2 250-2 257

[13] Klein G, Murray D. Improving the Agility of Keyframe-based SLAM[C]//Proceedings of European Conference on Computer Vision, 2008:802-815