[文章编号] 1003－4684(2025)02-0011-06

基于分区融合特征点提取的车载视频稳像方法

刘城逍[1]，吴春林[1]，姚思威[2]，何飞宇[3]

(1 湖北工业大学机械工程学院，湖北 武汉 430068；
2 湖北工业大学底特律绿色工业学院，湖北 武汉 430068；
3 湖北省产品质量监督检验研究院，湖北 鄂州 436000)

[摘 要] 车载视频在动态背景下往往会出现视频图像抖动的问题，而常用的 Harris 特征点提取算法在处理此类图像时，所提取的特征点常存在分布不均匀、易产生伪特征点等缺陷，进而影响电子稳像效果，由此提出一种基于分区融合特征点提取的稳像算法。首先利用 Prewitt 边缘检测提取图像的边缘信息，并对车载视频的边缘提取图像进行分区处理；再使用 Harris 和 AKAZE 算法对不同区域图像进行特征点提取；在完成分区特征匹配后，通过计算运动参数并进行加权融合，得到全局运动参数，使用卡尔曼滤波器平滑运动参数；最后，根据得到的运动参数对车载视频进行补偿，从而得到稳像后的视频。实验结果表明，改进后的算法显著提升了稳像效果，稳像后视频的 PSNR 值约提升了 5.2%，且具有较好的实时性。

[关键词] 电子图像稳定；Harris 特征提取算法；AKAZE 特征检测算法；边缘检测技术；特征点提取
[中图分类号] TP391.41 [文献标识码] A

现代智能汽车对车载视频系统提出了更高的行人跟踪等检测技术需求，而稳定的车载视频是确保识别精度的基础。然而，车辆在行进过程中，车载视频往往会出现高频小范围抖动，这严重降低了识别技术的检测精度。因此，需要一种成本低、对高频小范围抖动稳像效果好的电子稳像方法来解决视频抖动问题。

电子稳像主要涵盖运动估计、运动滤波和运动补偿等三个核心步骤。在运动估计方法中，基于特征点匹配法的运动估计相比于块匹配法[1]、灰度投影法[2]等有更好的实时性和稳定性[3]。基于特征点匹配法的运动估计算法可进一步细分为特征点提取、特征点匹配和全局运动估计三个环节，高效地提取特征点能显著提升后续特征点提取、特征点匹配和全局运动估计的效率。

目前，特征点提取算法种类繁多，其中 Harris[4]、SIFT[5]、Shi-Tomasi[6] 等算法较为常见。Harris 算法以其提取的特征点较为稳定、算法复杂度低、速度较快以及具有一定的旋转不变性等优点，较为适合车载视频的稳像。然而，车载视频在车辆行进中采集的图像背景是动态的，包含多个对象，且纹理和光照变化多样，这导致 Harris 算法在提取特征点时容易出现伪特征点和分布不均匀的问题，进而影响视频的稳像效果。

针对这一问题，张见双等[7]提出将图片分割成 9 份子图像，并根据每幅子图的对比度来计算每个子图 Harris 的阈值，从而减小了特征点聚簇现象，缺点是特征点检测时间有所增加；Rohit 等[8]使用 K 均值聚类算法对图像进行基于颜色的分割，再使用 Harris 算法检测特征点，并引入 FREAK(fast retina keypoint)描述符提高特征点匹配精度，但该方法对噪声较为敏感，易产生伪特征点；卢伟家等[9]提出了一种基于像素灰度值的加权均值滤波方法，通过实现自适应加权系数的局部移动窗口来有效抑制噪声对 Harris 算法的影响，但该方法对整幅图进行提取，计算量大，且实时性不佳。

针对单一特征点提取算法在车载视频图像特征点提取上存在效率和准确性不足的问题，本文提出一种基于分区融合特征点提取的车载视频稳像算法。该算法首先对车载视频图像进行边缘信息提取，并对提取后的图像进行分区处理，再采用 Harris 和 AKAZE[5] 特征点提取算法对各分区进行特征点提取，有效提升了特征点的提取速度和抗噪性。然后，通过融合不同分区特征点计算出的运动估计参数，提高了全局运动估计准确性。最后，运用卡尔曼滤波对运动参数进行处理，有效去除了随机抖动，最

[收稿日期] 2023－08－29
[基金项目] 国家自然科学基金(51575164)
[第一作者] 刘城逍(1994－)，男，湖北京山人，湖北工业大学硕士研究生，研究方向为机器视觉，嵌入式图像处理。

终实现了车载视频的稳像效果。

1 基于分区融合特征点提取的运动估计

1.1 基于边缘检测的分区特征点提取算法

Harris 特征点提取算法在处理动态背景、多样纹理和光照条件以及存在噪声的车载视频图像时，难以设定一个精确的阈值来同时满足特征点的均匀分布和提取效率。此外，Harris 算法对局部物的尺度变换较为敏感，难以稳定地提取出相同的特征点[10]。相比之下，AKAZE 算法对局部物体的尺度变换和旋转变换具有较好的鲁棒性，且运算速度可以基本满足车载视频稳像的需求。

因此，本文提出了一种分区融合特征点提取算法。该算法首先针对车载视频图像的特点，利用 Prewitt 边缘检测技术提取出原始图像的边缘特征。然后，基于车载视频图像中不同区域特性的差异，将图像划分为 A、B 两块区域。在这两块区域中，分别采用 Harris 和 AKAZE 算法进行特征点提取，最后合并两种算法的提取结果，以获得更为全面和准确的特征点信息。算法的流程如图 1 所示。

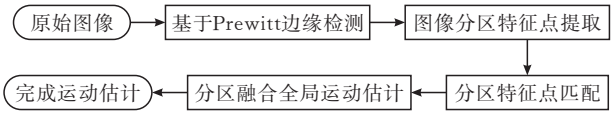


图 1 运动估计流程

1.1.1 基于车载视频图像特性的分区特征点提取

通过分析车载视频图像可知，图像中心区域背景复杂，且局部运动物体的尺度变换较为显著。同时，连续几帧图像的中心区域特征变化不大，而边缘区域呈现出明显的变化。中心区域所提取的特征点在前后帧中均存在，对车载视频稳像的效果具有更为重要的影响。而在图像的边缘区域，特别是车身部分，其几何结构清晰明了，特征简单且明显，属于刚性物体，且不存在尺度变换的问题。

针对两种算法的特性以及车载视频图像的独特性质，结合人类视觉对图像非均匀敏感的特性[11]，对图像进行了分区处理，并在此基础上进行特征点提取。具体而言，在图像（图 2）的中心区域选取一个矩形作为关注的核心区域，记为 A 区。在 A 区内，采用 AKAZE 特征点提取算法，以提高人眼注意力集中区域的特征点提取准确性，同时避免对全图进行提取而减慢提取速度。

此外，选取图像中车身一侧的区域作为边缘区域，记为 B 区，并在该区域内使用 Harris 特征点提取算法。根据生物非均质采样模型[11]的指导，确定中心区域 A 长宽为图像长宽的 3/5，而边缘区域 B 的宽度则为整幅图像的 1/5。



图 2 分区示意图

1.1.2 基于 Prewitt 算法的边缘检测

针对 Harris 和 AKAZE 算法在噪声影响下容易提取大量伪特征点的问题，采用边缘提取算法来优化特征点提取过程。相较于传统的高斯滤波、均值滤波等降噪算法，边缘提取算法能够更有效地凸显图像中物体的边界信息。这有助于特征点提取算法更准确地定位到物体边界附近的特征点，从而提升车载视频稳像的效率。

在边缘检测领域，Sobel、Prewitt、Canny 算法[12]是几种较为典型且广泛应用的方法。其中，Prewitt 算法以其出色的抗噪性能、清晰的线条表现以及较快的计算速度而著称。因此，本文选取了 Prewitt 算法来进行边缘检测。通过该算法，成功提取出了原始图像中的边缘信息，并保存了这些信息以用于后期的特征点提取。这一做法有效避免了噪声对特征点提取过程的干扰，显著减少了提取出的伪特征点数量。

1.2 分区特征点匹配

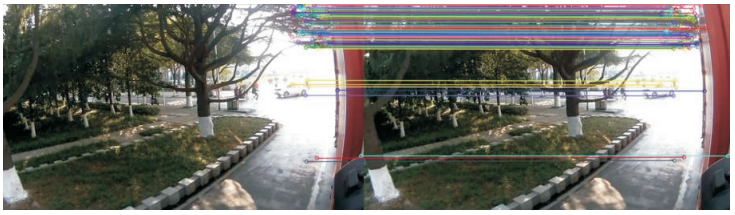
特征点的匹配过程是将两帧图像中提取出的特征点进行一一配对，其准确率直接关系到稳像的最终效果。针对本文中从两个区域提取出的特征点，分别采用 FLANN（快速最近邻搜索算法）[13]和 LK 光流法（Lucas-Kanade 光流法）[14]对 A、B 区域的特征点进行初步匹配，再利用 PROSAC 算法（渐进样本一致性算法）[15]从特征点集中随机抽取一小部分作为初始抽样集，并基于这个初始抽样集合计算模型，来计算每个特征点与当前模型的拟合度，并根据拟合度对特征点进行排序。通过这一过程，进一步筛选出高质量的特征点对，同时去除误匹配点。这一方法有效降低了算法的计算时间，提高了特征点匹配的准确率，从而提升稳像效果。如图 3 所示。

1.3 分区融合全局运动估计

完成图像特征点的匹配后，需要结合运动变化模型计算出图像的运动参数，本文分别计算 A 区域和 B 区域的运动参数，对不同区域的全局运动参数赋予不同权值，融合得到更准确的运动估计参数。以区域 A 为例，根据仿射变换原理，满足：



(a)B 区域初匹配结果



(b)B 区域 PROSAC 法去除误匹配结果

图 3　特征点匹配效果

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \boldsymbol{H} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & f_1 \\ a_3 & a_4 & f_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \tag{1}$$

式中：(x_1, y_1) 为参考帧特征点坐标；(x_2, y_2) 为与参考帧特征点匹配的相邻帧特征点坐标；$\boldsymbol{H}$ 为 A 分区中图像间相邻帧对应的仿射变换矩阵；a_1、a_2、a_3、a_4 为相邻帧的旋转参数；f_1、f_2 则表示 A 区间中相邻帧的水平和垂直运动。

A 区域的运动参数矩阵记为 $\boldsymbol{M}_1$，其对应的权值为 k_1；B 区域的运动参数矩阵记为 $\boldsymbol{M}_2$，其对应的权值为 k_2。则融合后的相机运动参数矩阵 $\boldsymbol{M}$ 应满足以下关系：

$$\begin{cases} \boldsymbol{M} = k_1 \boldsymbol{M}_1 + k_2 \boldsymbol{M}_2 \\ k_1 + k_2 = 1 \end{cases} \tag{2}$$

考虑到人眼对运动信息具有更高的敏感性，而车载视频图像的中心区域恰恰包含了大量的运动信息，因此，本文采用固定权值 $k_1 = 0.8$，$k_2 = 0.2$，这样既能保留中心区域数据的重要性，又能将边缘区域的参数作为补充，以实现更准确的相机运动参数估计。

2　运动滤波和运动补偿

运动滤波旨在减轻或消除运动模糊和抖动现象，同时保留相机的主动运动变化。在众多的运动滤波算法中，卡尔曼滤波[16]凭借其能够实时地结合前一帧运动参数的最优预测值与当前帧运动参数的测量值，从而得出最优预测结果的优势，特别适用于对实时性要求极高的车载视频稳像领域。

2.1　卡尔曼滤波

在车辆行进过程中，采集视频的相机所表现的主动运动呈现出低频特性，而其被动的随机抖动则展现出高频运动特征。考虑到这两种运动的差异，采用卡尔曼滤波对相机的运动轨迹进行精细化处理。基于卡尔曼滤波的运动平滑预测过程为：

$$\begin{cases} \hat{x}_k = A\hat{x}_{k-1} + B\,u_{k-1} + \omega_{k-1} \\ P_k = A\,P_{k-1}\,A^{\mathrm{T}} + Q \end{cases} \tag{3}$$

式中：$\hat{x}_k$ 为第 k 帧对应的先验预估值；$\hat{x}_{k-1}$ 为 $k-1$ 时刻预测值；A、B 为系统参数；u_{k-1} 为系统输入量；ω_{k-1} 为系统噪声偏差；P_k 为第 k 帧对应的先验估计误差协方差；P_{k-1} 为上一过程最优估计协方差；Q 为系统噪声的协方差。卡尔曼滤波的更新过程为：

$$\begin{cases} G_k = P_k\,H^{\mathrm{T}}\,(H\,P_k\,H^{\mathrm{T}} + R)^{-1} \\ \hat{x}_k = \hat{x}_k + G_k(z_k - H\hat{x}_k) \\ P_k = (I - G_k H)\,P_k \end{cases} \tag{4}$$

式中：G_k 为第 k 帧卡尔曼增益；H 为测量系统参数；R 为测量噪声的协方差；z_k 是 k 帧测量值；$\boldsymbol{I}$ 为单位矩阵。当 $Q = 0.01$，$R = 1$ 时，平滑效果最好。

2.2　运动补偿

运动补偿是电子稳像技术中的核心环节之一。在通过卡尔曼滤波获得平滑的运动矢量后，选取前一帧作为参考基准，运用双线插值法对相邻帧中的随机抖动进行精准补偿。具体做法是，在目标位置周边的四个最近邻像素点中进行插值计算，并依据各像素点距离目标位置的远近，对这些像素的灰度值进行加权平均。这一处理过程在图像上创建了平滑的过渡效果，有效避免了锯齿状边缘和块状效应的出现，使得参考帧与相邻帧的图像内容更加接近一致。最终，经过这一系列处理，获得稳像后的高质量视频。

3　实验与分析

为检验基于分区融合特征点提取算法的车载视频稳像技术的实际效果，依托湖北工业大学机械工

程学院所研制的车载电子视镜摄像系统进行实验。实验所采用的 CPU 为八核 Cortex-A53 处理器，操作系统为 Ubuntu16.04，视频分辨率设置为 640×480，帧率为 60 FPS。

3.1 特征点提取实验

在车载视频图像的实验中，采用 Harris 算法与本文提出的分区融合算法进行对比实验，实验结果如图 4 所示。其中，图 4a 展示了使用 Harris 算法提取的特征点结果，图 4b 则呈现了本文算法的特征点提取结果。通过对比可以看出，Harris 算法提取的特征点数量众多，且大部分特征点密集地聚集在图像上半部分的某一物体上。同时图像边缘部分的特征点也较多，这导致特征点之间难以区分，增加了误匹配的风险。而本文改进的算法，在设定相同阈值的条件下，提取的特征点分布更加均匀，未出现特征点大量聚集的现象。同时，该算法对边缘的平坦区域进行了有效过滤，未提取特征点，从而极大地提高了算法的效率。



(a)Harris 算法提取结果



(b)本文算法提取结果

图 4 特征点提取对比

为了评估特征点分布的均匀性，先遍历每一个特征点周围 5 个像素点范围内存在特征点的数量，再计算这些数量的平均值，以实现客观、量化的分析。若某个特征点周围 5 个像素范围内特征点的数目较多，则表明该区域的特征点分布相对不均匀。通过对一段 300 帧的车载视频进行特征点提取与统计，得到了具体结果(表 1)。结果中所有数据均保留至小数点后两位。

表 1 特征点分布对比

特征点提取方法	周围特征点个数
Harris	23.04
本文算法	2.66

由表 1 可知，经过本文算法提取的图像中，一个特征点周围 5 个像素点范围内平均仅存在 2.66 个特征点。这一数值相较于 Harris 算法下降了 88.63%，体现了本文算法在提取特征点分布均匀性上的优势。

3.2 稳像效果实验

3.2.1 稳像效果客观分析 在视频稳像效果的客观评价中，峰值信噪比(peak signal-to-noise ratio，PSNR)是一个常用的效果指标[17]。该指标通过比较两个信号之间的均方误差，来量化两幅不同图像之间的相似程度，PSNR 值越大，表明二者越相似，从而反映稳像的效果越好。

为了全面评估稳像效果，对稳像前后的两段视频逐帧进行分析。具体做法是：计算每一帧与相邻帧之间的 PSNR 值，并对这些 PSNR 值求取平均值。结果显示：稳像前的视频平均 PSNR 值为 31.66 dB，而稳像后的视频平均 PSNR 值则提升为 33.31 dB。相较于原视频，稳像后的视频平均 PSNR 值提高了 5.2%，这一数据充分表明稳像效果的显著改善。图 5 展示了稳像前后视频帧序列间 PSNR 值的对比情况。

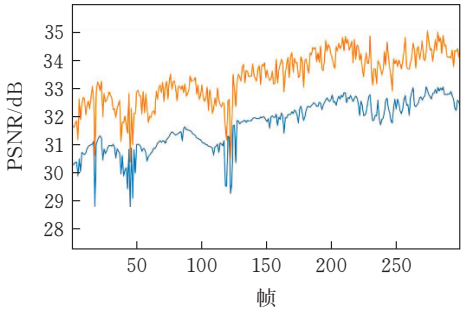


图 5 稳像前后的 PSNR 对比

3.2.2 稳像效果主观分析 为了直观展示效果差异，选取了稳像前后的相邻帧画面进行对比(图 6)。其中，图 6a 展示了稳像前相邻帧的画面。红色直线作为参考线，用于更加直接地评估稳像效果。

通过观察图 6 中路牌与参考线的距离对比，可以明显看出，在稳像前，由于抖动导致路牌向上偏移明显。而经过稳像处理之后，路牌向上偏移得到明显改善(图 6b)。



(a)稳像前的相邻两帧图像



(b)本文算法稳像后的相邻两帧图像

图 6　稳像前后的相邻帧对比

3.2.3　稳像算法实验结果对比　本文算法与其他几种稳像方法处理同一段视频的效果进行对比。结果发现，相较于基于 Harris 的稳像算法，其他三种算法在稳像效果上都有明显提升。Bansal 等[18]通过使用方形滤波器和滑动平均窗口对基于 ORB 算法的视频稳像进行了改良，Verma 等[19]采用基于改进后 FAST 算法的视频稳像方法，这两种方法都能够获得 PSNR 值更高的稳像视频，但在实时性方面却逊色于本文算法。本文算法的单帧处理时间为 32.25 ms，处理速度能达到 31 FPS，同时在 PSNR 指标上也取得了良好的结果。各算法的处理速度对比详见表 2。

表 2　稳像效果与速度

稳像方法	分辨率/P	PSNR/dB	单帧处理时间/ms
原视频	480	31.66	—
Harris	480	32.01	52.63
Bansal 等	480	34.08	76.92
Verma 等	480	35.45	58.82
本文算法	480	33.31	32.25

4　结论

针对车载图像的特点以及动态背景下特征点难以准确提取的问题，本文提出了一种车载视频稳像方法。该方法基于分区融合理念进行特征点提取。通过在车载视频上进行实验验证，得到如下结论。

1)利用边缘检测技术提取图像的边缘信息，并结合车载视频图像的特点和生物非均质采样模型，对图片进行分区处理。在分区后，采用 Harris 和 AKAZE 特征点提取方法进行特征提取，有效降低了噪声的干扰，并获得分布均匀的特征点。

2)在特征点初匹配后，运用 PROSAC 算法剔除误匹配点，同时对不同区域的运动估计参数赋予合理的权值并进行融合，显著提高了全局运动估计的精度。

3)利用卡尔曼滤波技术获取主动运动参数，并通过运动补偿处理，最终得到稳像后的车载视频。

4)实验对比分析结果表明，本算法在车载视频电子稳像方面表现出较好的性能。与原视频相比，稳像后的视频 PSNR 值达到 33.31 dB，提升了 5.2%，且处理速度基本满足实时性的要求。

[参　考　文　献]

[1] BHASKAR H，BEZ H E. A novel block motion estimation model for video stabilization applications[C]// Informatics in Control Automation and Robotics Robotics and Automation，2007：303-306.

[2] 范叶平，郭政，张锐.基于下采样灰度投影的电子稳像算法研究[J].工矿自动化，2017，43(04)：22-27.

[3] GUILLUY W，OUDRE L，BEGHDADI A. Video stabilization：Overview，challenges and perspectives[J]. Signal Processing：Image Communication，2021，90：116015.

[4] 王杰，经俊森，陈正伟，等.基于 Harris 和卡尔曼滤波的农业机器人田间稳像算法[J].农业机械学报，2023，54(01)：30-36.

[5] 宗慧琳，袁希平，甘淑，等.改进 AKAZE 算法的泥石流区无人机影像特征匹配[J].测绘通报，2023(02)：91-96.

[6] 彭晓星，罗霄，韩宝玲.一种改进的 Shi-Tomasi 角点检测方法[J].计算机应用与软件，2022，39(12)：241-245.

[7] 张见双，张红民，罗永涛，等.一种改进的 Harris 特征点检测的图像配准方法[J].激光与红外，2017，47(02)：230-233.

[8] TIWARI R K，VERMA G K A . A computer vision based framework for visual gun detection using harris interest point detector[J]. Procedia Computer Science，2015，54 ：703-712.

[9] 卢伟家，刘缠牢.一种基于 Harris 特征点检测的改进算法[J].仪表技术与传感器，2017(12)：98-100.

[10] 王正家，吴春林，柯黎明，等.基于 Harris 图像拼接的全景视频稳像算法[J].红外技术，2022，44(11)：1203-1209.

[11] 张斌，梅魁志，郭青.仿生物视觉的非均匀采样方法及其硬件设计[J].西安交通大学学报，2010，44(06)：99-103.

[12] QI J，YANG H T. Research on image segmentation and edge detection technology based on computer vision[C]// Journal of Physics：Conference Series. IOP Publishing，2021，1994(01)：012035.

[13] HUANG S，SUN G，LI M. FAST and FLANN for feature matching based on SURF[C]// 33rd Chinese Control and Decision Conference (CCDC). IEEE，2021：1584-1589.

[14] AHN B, HAN Y, KWEON I S. Real-time facial landmarks tracking using active shape model and lk optical flow[C]// 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). IEEE, 2012: 541-543.

[15] CHUM O, MATAS J. Matching with PROSAC-progressive sample consensus[C]// IEEE computer society conference on computer vision and pattern recognition (CVPR'05). IEEE, 2005, 1: 220-226.

[16] 李成城,马立森,田原,等.基于 CLAHE 与卡尔曼滤波的掘进机机载视频稳像算法[J].工矿自动化,2023,49(05):66-73.

[17] 尹丽华,李范鸣,刘士建.基于区域分割与融合的全景稳像算法[J].红外与激光工程,2018,47(09):391-400.

[18] BANSAL P, DINESH J B, VR S K, et al. Video stabilization using ORB detector[C]// 14th International Conference on Computer and Automation Engineering (ICCAE). IEEE, 2022: 50-55.

[19] VERMA K, SINGH R K, KUMAR B, et al. FAST based video stabilization with preserved intentional motion and smear removal[C]// International Conference on Electrical and Electronics Engineering (ICE3). IEEE, 2020: 376-381.

Vehicle Video Image Stabilization Method Based on Partition Fusion Feature Point Extraction

LIU Chengxiao[1], WU Chunlin[1], YAO Siwei[2], HE Feiyu[3]

(1 *School of Mechanical Engineering, Hubei Univ. of Tech., Wuhan* 430068, *China*;
2 *Detroit Green Industry College, Hubei Univ. of Tech., Wuhan* 430068, *China*;
3 *Hubei Institute of Product Quality Supervision and Inspection, Ezhou* 436000, *China*)

Abstract: In vehicle videos, there used to be video image jitter in dynamic backgrounds, and the commonly used Harris feature point extraction algorithm, when applied to extract feature points from such images, exhibited uneven distribution and produced pseudo-feature points. This affects the effectiveness of electronic stabilization. Consequently, this paper introduced a stabilization algorithm based on partitioned fusion feature point extraction. Initially, edge information was extracted from the images using Prewitt edge detection, and then in vehicle video frames were partitioned based on their edge content. Subsequently, the Harris and AKAZE algorithms were applied to extract feature points from different regions within the images. After completing the partitioned feature matching, motion parameters were computed and weighted, resulting in a fused global motion parameter. A Kalman filter was then utilized to smooth these motion parameters. Finally, compensation was applied to the in vehicle video frames to obtain stabilized video output, Experimental results demonstrate that the improved algorithm effectively enhances the stabilization performance, with an approximately 5.2% increase in the PSNR value of the stabilized video. Furthermore, it exhibits good real—time performance.

Keywords: electronic image stabilization; Harris; AKAZE; edge detection; feature points

[责任编校: 张 众]