

基于运动路径全曲率的视频稳定质量评价

郑清卓 张 磊 黄 华

(北京理工大学计算机学院 北京 100081)

摘 要 现有的无参考视频稳定质量评价算法存在对剧烈的抖动敏感或与主观评价结果一致性不高等问题. 针对这些问题, 提出了一种直接度量视频运动路径平滑程度的无参考视频稳定评价算法. 首先检测相邻两帧图像的特征点, 计算相邻帧间单应变换; 然后将单应变换映射到李群空间形成运动路径; 最后借助离散测地逼近方法计算运动路径全曲率, 以此衡量运动路径平滑程度, 进而评价视频稳定程度. 为了验证评价的有效性, 根据已公开的数据集构建了 150 段抖动视频及其稳定处理后的视频组成的数据集进行测试. 实验结果表明, 该文算法获得的视频稳定评价结果与人眼主观评价结果的相关性达到 97%, 相比于帧间保真度、频域分析等现有评价算法提高 20% 以上. 与全参考视频稳定质量评价算法的一致性对比结果进一步验证了该文算法的有效性.

关键词 视频稳定质量评价; 全曲率; 单应变换; 运动路径; 李群; 视觉运动感知

中图法分类号 TP391 **DOI 号** 10.11897/SP.J.1016.2018.02524

Video Stabilization Quality Assessment Based on Total Curvature of Motion Path

ZHENG Qing-Zhuo ZHANG Lei HUANG Hua

(School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081)

Abstract A video captured with a hand-held device (e.g., cell-phone or tablet computer) often appears remarkably shaky. It has a great negative impact on people's visual perception and is difficult for the following processing of the video. As a technology that can remove video jitter to enhance visual quality, video stabilization algorithms have been widely and deeply studied in the last ten years. However, how to effectively evaluate the performance of video stabilization algorithms is a topic that has not been studied deeply. Due to the lack of a comprehensive and fair evaluation mechanism, it is impossible to judge the performance of video stabilization algorithm objectively, which makes it difficult for users to choose video stabilization algorithm. The video stabilization objective assessment can be divided into two ways: non-reference and full-reference. The full-reference method needs the benchmark video, then calculates the difference between the evaluated video and the benchmark video, and then evaluates the stability of the video. Non-reference method is used to analyze the jitter characteristics of video directly through the motion information of video, and then evaluates the video stability. Compared with the full-reference methods, non-reference methods are more flexible and economical. But the existing non-reference video quality assessment algorithms have some problems, such as sensitive to video with severe jitter or inconsistent with the subjective assessment. In view of these problems, a non-reference video stability assessment method for directly measuring the smoothness of video motion path is proposed. Firstly, we detect the feature points of adjacent frames, and then calculate the inter-

frame homography matrix. Homography matrix can accurately describe the translation, rotation and other transformations between video frames. Secondly, we map the homography matrices in the Lie Group space to form the motion path. Finally, we calculate the total curvature by using the discrete geodesic approximation method to measure the smoothness of the motion path, and use it to evaluate the stability of the video. In order to verify the efficacy of the evaluation, we construct a dataset of 150 video clips and their corresponding stabilized videos based on the publicly available dataset. Stabilized videos in this dataset are the results of five video stabilization algorithms. We choose 2 classic stabilization algorithms from published literature: the spatially and temporally optimized method (STO) and bundled paths methods (BP). Besides, we choose three sophisticated softwares that realise the stabilizer function: Adobe After Effects (AE) warp stabilizer, Google YouTube stabilizer and VirtualDub Deshaker. The experimental result shows that the correlation between the results obtained by the proposed algorithm and the subjective evaluation results is 97%. We make comparisons with some other classic video stabilization quality assessment methods based on this dataset. We collect two representative methods that are used in previous video stabilization methods: inter-frame transformation fidelity (ITF) and low/high-frequency rate (LHR). Compared with the evaluation algorithm based on ITF, the correlation improves 39%. Compared with the evaluation algorithm based on LHR, the correlation improves 21%. The consistency with the full-reference video stabilization quality evaluation algorithm further verifies the effectiveness of the proposed algorithm.

Keywords video stabilization quality assessment; total curvature; homography; motion path; Lie Group; visual motion perception

1 引言

随着视频拍摄设备的廉价化和大众化,视频拍摄变得越来越容易,甚至催生了“全民直播”时代。但是,由于使用环境和和使用人员的限制,运动环境下拍摄的视频通常存在不同程度的画面抖动。例如,监控摄像头在大风环境下拍摄视频或是业余爱好者在没有三脚架、云台等专业的摄像稳定设备的情况下拍摄视频。视频稳定作为一种能够去除视频抖动以增强视觉质量的技术,被广泛而深入的研究。常用的视频稳定方法主要包括三种:2D的方法、3D的方法和2.5D的方法。2D方法以基于L1优化的视频稳定^[1]为代表,使用矩阵描述摄像机的运动轨迹,并对变换矩阵进行平滑,获得稳定的视频数据。3D方法以基于内容保持的3D视频稳定^[2]为代表,主要依赖于运动恢复结构(SfM)方法进行3D场景和相机运动轨迹的重建,根据重建结果获得稳定后的视频。基于2.5D的方法以子空间稳定^[3]为代表,使用特征点轨迹描述摄像机运动,对于轨迹求取其子空间,在子空间上对轨迹进行平滑,进而恢复得到全部的稳定视

频轨迹。最近两年也有许多优秀的视频稳定算法被提出,例如Zhang等人提出的一种快速视频稳定算法^[4]。然而,如何有效地评价视频稳定算法的性能,是一个没有被深入研究的问题。由于缺少一个全面公平的评价机制,无法对视频稳定算法进行客观的性能评判,导致用户很难选择合适的视频稳定算法。

与视频质量评价类似,视频稳定质量评价也可以分为主观评价和客观评价两种。主观评价由于需要大量的人力,不能适应大规模数据集的质量评价,一般用于评估客观评价的性能。客观评价分为无参考和全参考两种方式。目前视频稳定算法的客观性能评测主要采用全参考的方法。全参考方法通过视频抖动合成算法,可以由真实的稳定视频生成抖动的视频,然后使用视频稳定算法对抖动的视频进行稳定处理,得到稳像后的视频与真实的稳定视频进行对比,就可以判断稳定算法的效果。全参考方法需要真实的稳定视频作为基准^[5-6],但是这种基准视频一般使用固定摄像头拍摄的方式或者高分辨率图像生成的方式获取,无法涵盖大部分使用场景。此外,合成的抖动很难模拟现实场景产生的抖动,所以生

成的抖动视频不够真实,而无参考视频稳定质量评价通过统计模型可以得到视频的抖动特征,以此判断稳定后视频的抖动程度,进而评价稳定算法的性能,这种方式灵活且经济.因此本文选择无参考方式对视频稳定算法进行客观评价.

当前,视频领域的无参考质量评价算法不断更新,例如 Wu 等人提出基于重构的无参考视频质量评价方法用于量化视频失真程度^[7].视频稳定性是视频质量好坏的衡量标准之一,但当前视频质量评价中一般只关注视频失真情况,并没有考虑视频稳定程度.一种常见的无参考视频稳定质量评价方法是通过用户调查来获取主观评判分数,以此作为评价标准.现有的视频稳定文献^[8]中一般所使用这种方式证明文中算法的有效性,但这种方法所需要的人力成本比较大.帧间峰值信噪比(PSNR)可以衡量两帧之间的相似性,能够作为一种度量直接评价视频的稳定质量^[9].但是这种方法对剧烈抖动的视频非常敏感.Liu 等人^[10]提出了一种在频率域上评价视频稳定性的方法,通过视频运动中的低频部分所占的比例衡量视频稳定质量.然而这种方法需要设置分隔高频和低频的阈值,通用性较差.

运动稳定性本质上是一种感知属性,需要探索合理的数学模型进行解释.从运动轨迹的角度来看,稳定视频的运动轨迹一般是平滑的,抖动视频的运动轨迹一般会具有较多的波峰和波谷^[10].而在数学上,曲率用于度量曲线的弯曲程度:平滑的曲线具有较小的曲率;有较多波峰波谷的曲线具有较大的曲率.于是,本文直接从运动平滑程度入手研究无参考稳定质量评价,提出一种基于运动路径全曲率的无参考视频稳定评价算法,利用运动路径的全曲率衡量视频运动平滑程度,进而评价视频稳定性.

本文的主要贡献在于提出一种新的无参考视频稳定质量评价算法,利用帧间变换形成的运动路径的全曲率作为度量,评价视频稳定算法效果的优劣.在已公开的抖动视频及其稳像视频构成的数据集上的测试,验证了本文算法的灵活性和高效性,并且与已有算法相比,具有更高的主观一致性.

本文在第 2 节介绍论文的相关研究工作;在第 3 节对论文算法进行简要介绍;在第 4 节详细介绍本文提出的基于运动路径全曲率的无参考视频稳定质量评价算法;在第 5 节通过大量实验验证本文算法的有效性;在第 6 节总结论文并展望未来的研究方向.

2 相关工作

最近十年大量的视频稳定算法被提出,其中有许多算法能够获得良好的稳定效果.例如 Grundmann 等人提出 L1 路径优化算法^[1],用 L1 范数约束平滑的运动轨迹的组成成分,通过求解一个线性规划问题来得到平滑后的轨迹.Liu 等人提出子空间分解算法^[3],先通过特征点跟踪得到很多较长的轨迹,然后对一段时间内的轨迹做子空间分解,得到特征轨迹,接着只对特征轨迹进行平滑处理,最后通过分解系数和平滑后的特征轨迹重构平滑的特征点轨迹.Liu 等人的捆绑路径优化算法^[10],首先对视频不同的区域分别计算局部运动轨迹,再对这些运动轨迹一起进行平滑处理,在处理过程中加入一些约束条件以保证平滑后轨迹的相似性.但文献中对这些算法的稳定效果评估主要依据人眼主观评价,缺乏一个统一的评价体系对这些算法进行客观的性能评判.

对视频稳定算法性能的评价主要是抖动的去除程度,与视频质量评价方法一样,分为主观评价和客观评价两个方面.主观评价是当前视频稳定文献中常用的评估方法,它主要基于人眼视觉系统.目前最具代表性的主观评价方法是主观意见分(Mean Opinion Score, MOS),即通过用户调查的方式对视频主观打分的平均值来表示视频质量的好坏.例如 Panetta 等人^[11]使用主观意见分验证图像质量评价算法的有效性.主观评价的主要问题是耗时多、需要进行多次重复实验且无法用数学模型描述,但它的评价结果能够真实反映视频的直观质量、稳定可靠,对于以提高用户观感舒适度为目的的视频稳定算法而言是必不可少的.所以本文在实验部分使用用户调查的方式获取主观意见分,进而证明本文算法与人眼主观评价具有一致性.

客观评价是设计一些能反映视频稳定质量的度量标准,通过具体的公式来计算这些度量的值,以此来反映视频稳定质量,一般分为全参考和无参考两种方式.全参考质量评价方法需要真实稳定的视频作为评价的参考,通过计算稳定视频和真实稳定视频的差异来度量稳定视频的质量.Offiah 等人^[12]提出对比稳定处理后的视频与真实稳定视频的图像质量,以此评估视频稳定算法的效果.Tanakian 等人^[5]提出计算稳定处理后的视频与真实稳定视频之间的相机运动轨迹差异,用来评估稳定算法的效果.

这两种评价指标计算简单、有明显的物理意义, 容易用于优化算法中, 因此应用广泛。但它们跟人眼视觉质量结果没有很好的一致性。无参考质量评价方法不需要真实稳定的视频, 直接通过统计模型得到视频的抖动特征, 以此判断稳定后视频的抖动程度, 进而评价视频稳定算法的性能。Battiatto 等人^[9]提出计算相邻帧间的峰值信噪比(PSNR)来反映稳定处理后视频的帧间保真度, 若处理后的视频比较稳定, 即路径比较平滑, 则相邻帧的位移和旋转都会较小, 计算出的 PSNR 值较大, 值越大稳定效果越好。这种方法对于一些抖动程度剧烈的视频无法达到很好的评价效果。Liu 等人^[10]在其工作中采用了裁剪率(cropping)和稳定性(stability)两个指标来评价稳定后视频的质量。裁剪率表示稳定后视频相对于抖动视频所保留内容的比率, 比率越大表明保留原始视频的信息越多。稳定性则是一种频域分析指标, 它表示稳定后轨迹的低频分量能量占总能量的百分比, 值越高表明低频分量越多, 轨迹越稳定。但这种稳定性衡量方式需要设定高频/低频的分隔阈值, 面对不同运动类型的视频, 这一阈值设置显然是不同的, 所以这种方法的通用性较差。

相比全参考质量评价方法, 无参考的评价方法不需要真实稳定的视频, 比较符合实际的需要。并且当前主流的视频稳定算法主要通过平滑轨迹的方式稳定视频。因此本文提出一种直接测量稳像视频运动路径平滑程度的无参考视频稳定评价算法。不同于以往的无参考视频稳定评价算法, 本文首先将视

频帧间单应变换映射到李群空间形成运动路径, 并在度量视频运动路径平滑程度时首次引入数学中的全曲率概念, 从而提高算法评价结果与人眼主观评价的一致性。大量实验也验证了本文提出的无参考视频稳定质量评价算法的有效性。

3 算法概述

本文利用运动路径的几何平滑性来衡量视频运动的平滑程度, 进而评价视频的稳定性。本文算法流程如图 1 所示, 主要包含如下 3 个部分:

(1) 特征点检测与匹配。采用加速鲁棒特征(SURF)算法^[13]检测相邻帧特征点。因为 SURF 算法采用图像点的二维微分 Hessian 矩阵^[14]的行列式来检测角点, 能很好地反映图像局部的结构信息; 再利用快速最近邻逼近搜索(FLANN)匹配特征点, 然后通过 K 最近邻(KNN)分类算法筛选匹配点, 这两种算法进一步提高了匹配精度。

(2) 李群空间中的运动路径建模。利用匹配成功的特征点计算相邻两帧的单应矩阵, 经过全部视频帧的计算, 将得到的所有单应矩阵组成运动路径。这条运动路径满足具有黎曼度量的李群空间的性质, 可以看作嵌入李群空间中的一条测地曲线。

(3) 全曲率计算。利用测地离散逼近方法首先计算运动路径全曲率。以全曲率为度量比较多种当前最流行的视频稳定算法。

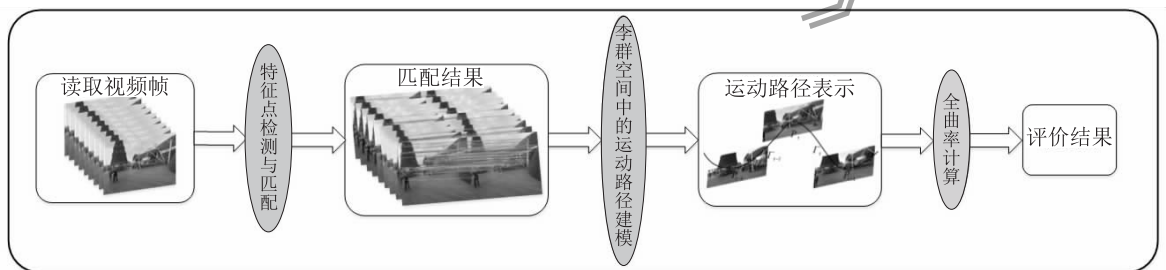


图 1 本文算法流程图

4 基于全曲率的稳定评价

为了获得视频运动路径, 首先需要进行特征点检测与匹配, 从而获得对应的特征点对, 然后计算两帧之间的单应矩阵。本文使用加速鲁棒特征算法检测图像中的特征点, 主要包含 5 步:

(1) 构建 Hessian 矩阵, 生成所有的兴趣点, 用

于特征提取。

- (2) 构造尺度空间。
- (3) 定位特征点。
- (4) 确定特征点主方向。
- (5) 构造特征描述子。

然后, 本文使用快速最近邻逼近搜索匹配相邻两帧的特征点, 并通过 K 最近邻分类算法筛选匹配点。最后, 通过匹配后的特征点对便可以计算出相邻

两帧的单应矩阵,从而构成运动路径.

本文的一个关键思想是利用运动路径的几何平滑性来衡量视频运动的平滑程度,其中运动路径的平滑程度可以由其全曲率表示.因此,需要提出一个运动路径模型,例如嵌入适当空间的一条曲线,这样就可以借助这条曲线的全曲率作为运动路径平滑程度的度量.

4.1 李群空间的运动路径

从一个视频帧序列的相邻帧间变化可以获得 2D 单应变换矩阵,而这种帧间单应变换可以作为表示运动路径的有效模型^[10].

通常,2D 单应变换被定义为一个平面到另一个平面的投影映射^[15].数学上,它可以表示为一个具有 8 个自由度且行列式为 1 的 3×3 实矩阵.给定连续两帧图像 f_i 和 f_{i+1} ,就可以计算得到一个单应矩阵 H_i ,利用这个单应矩阵就可以将 f_i 和 f_{i+1} 对齐.然后,运动路径可以被表示为连续帧间单应矩阵的乘积,例如 $P(t) = P(t-1)H_t$,显然 $P(t)$ 依然是单应矩阵并且包含相应的运动信息.

从代数表示以及微分几何的观点来看^[16],所有的 2D 单应矩阵本质上都属于一个李群 \mathcal{H} ,称为特殊线性群:

$$\mathcal{H} = \{H \in R^{3 \times 3} \mid \det(H) = 1\} \quad (1)$$

其中 $\det(\cdot)$ 表示矩阵的行列式.

帧间单应变换映射到李群空间便形成运动路径曲线(如图 2 所示).

由于李群具有流形结构,它支持群内元素的微分和积分运算.特别地,属于李群 \mathcal{H} 的元素 H 的切线空间是一个矩阵向量空间,这个空间中矩阵的迹均为 0,即 $T_H = \{\hat{H} \in R^{3 \times 3} \mid \text{tr}(\hat{H}) = 0\}$.进一步,根据李群空间中的性质可以定义流形 \mathcal{H} 在向量空间中基

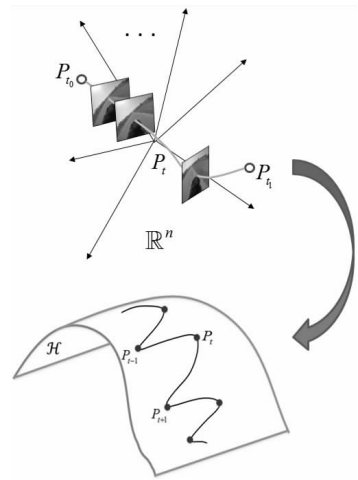


图 2 帧间单应变换形成运动路径

于内积的度量,表示为 $\langle \hat{H}_1, \hat{H}_2 \rangle_{T_H} = \text{tr}(\hat{H}_1^T \hat{H}_2)$,其中 $\text{tr}(\cdot)$ 表示矩阵的迹.

4.2 基于路径全曲率的运动平滑度量

运动路径 $P(t)$ 是李群 \mathcal{H} 中的一条曲线,它也是嵌入欧式空间 $R^{3 \times 3}$ 中的一个平滑子流形.由数学中曲线的定义可知,曲线的弯曲程度与它的曲率有关,而运动路径这条曲线的弯曲程度又可以衡量运动的平滑程度.同样从运动感知的观点看来,平滑的运动一般会表现出良好的曲率特性^[17].基于上述两点理论,我们将运动路径全曲率作为评价视频稳定质量的度量,如图 3 所示,图 3(a) 中的 3 条轨迹是由 3 种视频稳定算法处理原始抖动视频后获得的稳像视频的运动轨迹.将这 3 部视频计算得到的帧间单应变换映射到李群空间就会形成对应的运动路径,如图 3(b) 所示.稳定性好的视频形成的运动路径较为平滑且曲率较小,稳定性不好的视频形成的轨迹会有较多波峰波谷,曲率也较大,如图 3(c) 所示.

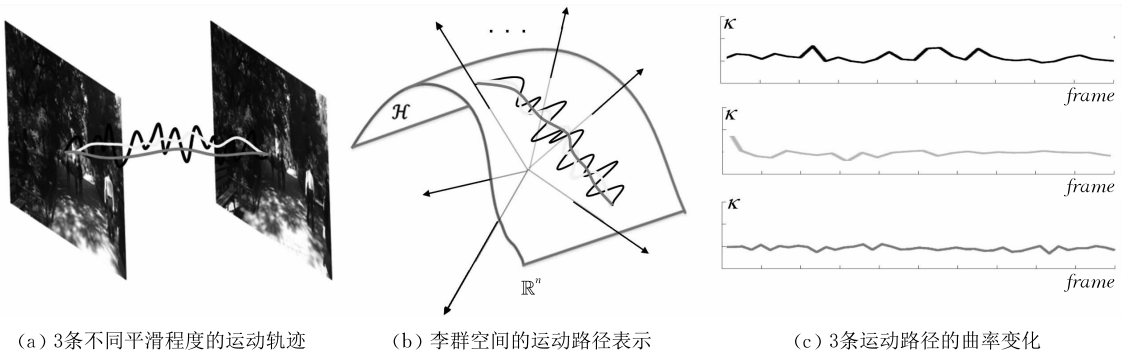


图 3 运动轨迹平滑程度与曲率的关系

数学上,设 $c(s)$ 为一条曲线,每一点的曲率为 $k(s)$,则积分 $\int_c k(s) ds$ 为 $c(s)$ 的全曲率.一条嵌入任

意黎曼流形中的曲线的全曲率可以看作这条曲线的内接测地多边形曲率之和的最小上界^[18].所以我们可以用测地多边形曲率之和来近似逼近曲线全曲率.令

\mathcal{M} 表示维度为 m 的黎曼流形, 其上存在一个测地多边形由一个测地线序列 $\Pi = (\Gamma_0, \dots, \Gamma_n)$ 组成, 测地线 Γ_i 的终点与 Γ_{i+1} 的起点是相等的, 测地多边形的顶点 x_0, \dots, x_n 表示这些测地线的起点和终点, 如图 4 所示, 其中黑色曲线表示黎曼空间中的曲线, 灰色曲线表示它的内接测地多边形。

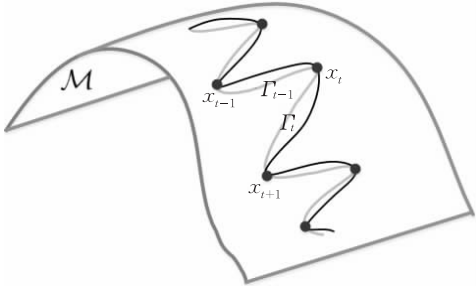


图 4 黎曼流形中的测地多边形

对于任意测地多边形 Π 都存在角度 α_i , 且有 $0 \leq \alpha_i \leq \pi$, 其中 $1 \leq i \leq n-1$. 这个角度表示切线向量 $\mathbf{T}_{i-1,t}$ 和 $\mathbf{T}_{i,t+1}$ 之间的夹角大小, 其中 $\mathbf{T}_{i-1,t}$ 表示测地线 Γ_{i-1} 终点的切向量, $\mathbf{T}_{i,t+1}$ 表示测地线 Γ_i 起点的切向量. 于是测地多边形 Π 的曲率和可通过如下公式计算:

$$\mathcal{K}(\Pi) = \sum_{i=1}^{n-1} \alpha_i \quad (2)$$

当 n 趋近于无穷时, 曲率和的极限就是这条曲线的全曲率, 因此式 (2) 是全曲率的近似逼近. 下一小节, 将具体介绍李群空间中基于单应变换的运动路径的全曲率计算, 并用它作为路径平滑程度的度量.

4.3 单点曲率计算

根据 3.1 节内容, 本文将单应变换矩阵形成的运动路径转化为黎曼空间中的一条曲线. 根据 3.2 节内容, 本文利用测地多边形的曲率和逼近这条曲线的全曲率. 通过这两步可以衡量这条路径的平滑程度, 进而评价视频稳定质量. 下面具体介绍测地多边形上单点曲率的计算方式, 然后沿曲线求和就可以求得测地多边形的曲率和.

如图 5 所示, 图中黑色曲线表示李群空间中的运动路径对应的内接测地多边形, 点 P_i 是测地多边形的其中一个顶点, 这一点的曲率可以由两个切向量之间的夹角 α_i 表示, 这两个切向量分别是 P_{i-1} 到 P_i 之间的测地线 Γ_{i-1} 的终点切向量 \mathbf{v}_{i-1} 以及 P_i 到 P_{i+1} 之间的测地线 Γ_i 的起点切向量 \mathbf{v}_i , P_i 被表示为连续帧间单应矩阵的乘积, 如下式:

$$\mathbf{P}_i = \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_i \quad (3)$$

其中, $\mathbf{H}_i: f_i \rightarrow f_{i+1}$ 是相邻两帧单应变换矩阵.

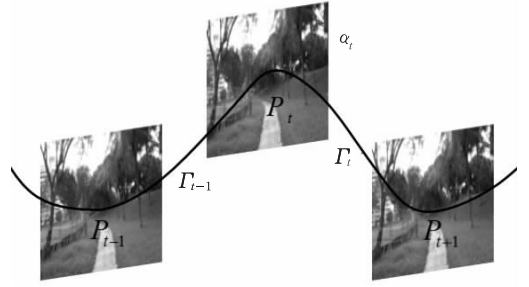


图 5 单应变换路径的曲率表示

对于第 t 帧, 计算该时刻单应变换路径的曲率首先需要计算从第 $t-1$ 帧到第 t 帧的测地线 Γ_{t-1} . 假设第 $t-1$ 帧到第 t 帧的单应变换为 \mathbf{H}_{t-1} , 那么根据文献[19]中的测地线概念, 这两帧之间的测地线 Γ_{t-1} 可以表示为

$$\gamma(\bar{i}) = \text{Exp}_{\mathbf{P}_d}(\bar{i} \text{Log}_{\mathbf{P}_d} \mathbf{H}_{t-1}), \quad \bar{i} \in [0, 1], \mathbf{H}_{t-1}: f_{t-1} \rightarrow f_t \quad (4)$$

其中, $\gamma(\bar{i})$ 表示第 $t-1$ 帧到第 t 帧的测地线 Γ_{t-1} , \bar{i} 表示测地线的位置, $\bar{i}=0$ 表示测地线起点, $\bar{i}=1$ 表示测地线终点. $\text{Exp}_{\mathbf{P}_d}(\cdot)$ 表示黎曼指数函数. $\text{Log}_{\mathbf{P}_d}(\cdot)$ 表示黎曼对数函数, 它是黎曼指数函数的逆映射. $\text{Exp}_{\mathbf{P}_d}(\cdot)$ 具体可通过如下公式计算:

$$\text{Exp}_{\mathbf{P}_d}(\mathbf{U}) = \exp(-\mathbf{U}^T) \exp(\mathbf{U} + \mathbf{U}^T) \quad (5)$$

其中, $\exp(\cdot)$ 表示矩阵指数函数. 根据矩阵指数求导公式,

$$\frac{d e^{t\mathbf{A}}}{dt} = \mathbf{A} e^{t\mathbf{A}} = e^{t\mathbf{A}} \mathbf{A} \quad (6)$$

可以计算测地线 Γ_{t-1} 的切向量如下式:

$$\begin{aligned} \frac{d}{d\bar{i}} \gamma(\bar{i}) &= \frac{d}{d\bar{i}} \exp(-\bar{i} \mathbf{U}^T) \exp(\bar{i} \mathbf{U} + \bar{i} \mathbf{U}^T) \\ &= ((-\mathbf{U}^T) \exp(-\bar{i} \mathbf{U}^T)) \exp(\bar{i} \mathbf{U} + \bar{i} \mathbf{U}^T) + \\ &\quad \exp(-\bar{i} \mathbf{U}^T) ((\mathbf{U} + \mathbf{U}^T) \exp(\bar{i} \mathbf{U} + \bar{i} \mathbf{U}^T)) \quad (7) \end{aligned}$$

令 $\bar{i}=1$, 可求出测地线 Γ_{t-1} 终点的切向量 \mathbf{v}_{t-1} , 表示如下:

$$\begin{aligned} \mathbf{v}_{t-1} &= ((-\mathbf{U}^T) \exp(-\mathbf{U}^T)) \exp(\mathbf{U} + \mathbf{U}^T) + \\ &\quad \exp(-\mathbf{U}^T) ((\mathbf{U} + \mathbf{U}^T) \exp(\mathbf{U} + \mathbf{U}^T)) \\ &= \exp(-\mathbf{U}^T) \mathbf{U} \exp(\mathbf{U} + \mathbf{U}^T) \quad (8) \end{aligned}$$

其中, $\mathbf{U} = \text{Log}_{\mathbf{P}_d} \mathbf{H}_{t-1}$. \mathbf{U} 可以通过软件 matlab 中的非线性最小二乘优化 $\text{lsqnonlin}(\cdot)$ 函数求解.

同理得到第 t 帧到第 $t+1$ 帧的测地线 Γ_t 起点的切向量 \mathbf{v}_t 如下式:

$$\mathbf{v}_t = \text{Log}_{\mathbf{P}_d} \mathbf{H}_t \quad (9)$$

由于两个切向量存在于流形 \mathcal{H} 的切线空间中, 根据 3.1 节的内容, 它们的内积可求, 于是本文使用下式计算夹角的余弦值:

$$\cos \alpha_i = \frac{\langle \mathbf{v}_{i-1}, \mathbf{v}_i \rangle}{\|\mathbf{v}_{i-1}\| \cdot \|\mathbf{v}_i\|} \quad (10)$$

其中, $\langle \mathbf{v}_{i-1}, \mathbf{v}_i \rangle = \text{tr}(\mathbf{v}_{i-1}^T \cdot \mathbf{v}_i)$ 且满足如下关系:

$$\|\mathbf{v}_{i-1}\| = \sqrt{\text{tr}(\mathbf{v}_{i-1}^T \cdot \mathbf{v}_{i-1})}, \|\mathbf{v}_i\| = \sqrt{\text{tr}(\mathbf{v}_i^T \cdot \mathbf{v}_i)} \quad (11)$$

通过上式计算得到这一点的曲率 α_i , 利用式(2)就可以得到测地多边形的曲率和, 然后本文将得到的曲率和除以 $n-1$ 从而获得这条运动路径的平均曲率值, 平均曲率值越小表示运动路径越平滑, 也就表示视频稳定算法处理后的视频越稳定.

5 实 验

为了证明本文提出的评价算法的有效性, 我们建立了一个数据集, 里面包含原始抖动视频以及利用各种视频稳定算法处理抖动视频得到的稳像视频. 实验主要包括 3 部分: (1) 将本文评价结果与人眼主观评测结果作对比, 利用对比结果证明本文评价算法与人主观评价具有一致性; (2) 将本文评价结果与其他视频稳定评价算法的结果比较, 利用对比结果证明本文算法具有优越性; 与视频质量评价算法作对比, 利用结果证明本文算法在稳定性度量方面具有优势; (3) 增加真实稳定的视频段进行测评, 并将本文评价结果与典型的全参考视频稳定质量评价算法结果比较, 进一步证明本文算法的正确性.

所有实验的测试环境为一台具有 3.30 GHz 英特尔双核处理器和 8GB 内存的台式机. 本文的稳定

评估算法运行时间大约为 1.6 秒每帧, 主要时间花费在 SURF 算法进行特征检测以及最终的曲率计算. 下面, 我们将详细介绍本文数据集以及实验结果.

5.1 数据集

本文数据集主要包含两部分. 第一部分是抖动视频以及用 5 种经典的视频稳定算法处理抖动视频得到的稳像视频. 原始抖动视频来源于 SIGGRAPH 2013 论文中已公开的数据集^[10], 共计 150 个视频. 我们将这些视频分为 7 类, 包括 crowd、driving、parallax、rotation、regular、running、zooming, 如图 6. 这些不同类别的视频涵盖不同的运动类型, 确保了数据集的多样性. 数据集集中的稳像视频部分由 5 种视频稳定算法处理抖动视频获得. 我们选择了两篇经典的视频稳定论文中的算法: 时空最优化算法(STO)^[20]和捆绑路径算法(BP)^[10]. 然后选择了三种完善的视频稳定软件: Adobe After Effects(AE)变形稳定器、Google YouTube 稳定器和 VirtualDub Deshaker. 事实上, AE 稳定器基于子空间算法^[3], 而 YouTube 稳定器主要实现了帧间变换上的一阶范数最优化^[1]. 因此, 数据集集中的每个抖动视频我们都可以获得对应的 5 种稳像视频, 它们表现出不同的稳定效果, 第一部分数据集共包含 900 个视频. 本文使用第一部分数据视频分析本文算法结果与人眼主观评价结果的一致性, 并与两种无参考视频稳定质量评价算法和一种无参考视频质量评价算法作对比.



图 6 数据集包含的 7 类视频

第二部分数据包含真实的稳定视频, 用于与典型的全参考视频稳定质量评价算法结果比较.

与图像质量评价一样, 由于人是视觉信号的最根本接受者, 视频稳定质量评价算法首先应该与人眼主观评测结果一致. 下一小节我们将证明本文评

价算法与人主观评价具有一致性.

5.2 结果与主观一致性检测

本文的稳定质量评价方式是首先根据第 3 节的评价算法计算原始视频以及它的 5 部稳像视频的运动路径平均曲率值, 然后根据平均曲率值的大小将这

6 部视频排序, 排序结果为 1~6, 1 表示曲率值最小、稳定性最好, 6 表示曲率值最大、稳定性最差. 本文采用在数据集上进行用户调查的方式获取主观质量评价结果, 以此作为标准评估我们的质量评价结果.

用户调查: 目的是基于观看视频这样的视觉实验, 让用户将 6 部视频按稳定程度排序. 本实验总共邀请了 50 个人进行主观评价并设计了一个软件协助用户观看及评价视频稳定性, 如图 7 所示. 每当用

户进行评价时, 软件从数据集中随机选取一组视频, 其中包含一个原始抖动视频以及 5 部稳像视频, 6 部视频在界面中的显示位置是随机的, 用户不知道视频所属分类也不知道稳像视频是由哪一种稳定算法处理. 用户可同时播放多个稳像视频和原始视频, 可以回放或快进甚至暂停在某一帧来帮助用户观看视频, 也可以单独播放一个视频来评估它的质量, 最终将 6 部视频按 1~6 排序并提交.



图 7 用户调查界面

每位志愿者完成对每组视频中的 6 部视频排序后, 每组视频中的每个视频都会被计算一个平均主观分, 然后按照平均分的大小再进行一次排序, 这样一个排序结果就可以表现 6 部视频的最终主观排序结果.

为了保证主观评价的可信度, 本文对主观评估的结果数据, 首先分析有无不合理数据或异常值的出现. 这里使用的是 95% 的置信区间, 置信区间之外的数据存在异常值的概率极高, 对数据整体的均值和标准差影响很大. 所以如果有异常值出现, 或有数据丢失现象, 则对相应的观察者进行剔除、修正数据或者补齐数据等处理, 才会进一步将其数据用于视频稳定质量方面的分析.

主观一致性: 视频稳定质量评价的有效性可以通过计算本文算法与用户调查的排序结果之间的一致程度来证明. 在统计学角度, 两组排序结果之间的相关性可以用斯皮尔曼相关系数衡量^[21]. 具体地, 假设两个已经排好序的随机变量分别为 $\{x_i\}$ 、 $\{y_i\}$, 元素个数均为 n , 则它们之间的斯皮尔曼相关系数被定义为下式:

$$r_s = 1 - \frac{6 \sum_{i=1}^n |x_i - y_i|^2}{n(n^2 - 1)} \quad (12)$$

显然, 两个排序变量相关性越高则斯皮尔曼相关系数越接近 1. 在本文的实验中, 计算数据集中每一类视频的平均稳定质量评价结果与用户调查的主观评价结果的斯皮尔曼相关系数. 结果如表 1 所示.

视频稳定算法评估: 从表 1 中的排序结果可以看出 5 种稳定算法都具有一定的稳定效果, 其中 STO 和 BP 算法的稳定效果最好, 因为它们在与空间上都进行了最优化的运动平滑. 但是由于 STO 算法依赖于帧间连续的轨迹检测, 当遇到剧烈抖动或者特征点不容易追踪的视频它的效果就会变差, 例如 rotation、running 类型的视频. BP 算法采用了划分网格的方式计算多条捆绑路径, 并分别进行运动平滑, 所以对 parallax 这类视频具有很好的处理效果. AE、Deshaker、YouTube 三种算法对不同的场景下的原始抖动视频都具有一定的稳像效果. 相比较 STO 算法, 它们具有较高的鲁棒性, 能够适应大多数的视频类型. 在部分场景下它们的稳像性能不如 STO 算法, 这主要是由于 STO 算法采用自由曲线表示平滑的运动路径, 所以稳定后视频的路径更加平滑.

表 1 本文评价算法得到的不同稳定算法的排序结果(括号中的排序是人眼主观评价结果)

视频类别	STO	AE	YouTube	Deshaker	BP	原始抖动视频	相关系数
crowd	1(1)	3(4)	4(3)	5(5)	2(2)	6(6)	0.94
driving	1(1)	4(3)	3(4)	5(5)	2(2)	6(6)	0.94
parallax	2(2)	4(4)	3(3)	5(5)	1(1)	6(6)	1.00
rotation	6(6)	3(3)	2(1)	4(4)	1(2)	5(5)	0.94
regular	1(1)	4(4)	3(3)	5(5)	2(2)	6(6)	1.00
running	3(3)	4(4)	2(2)	5(5)	1(1)	6(6)	1.00
zooming	1(1)	4(4)	3(3)	5(5)	2(2)	6(6)	1.00
平均	2(2)	4(4)	3(3)	5(5)	1(1)	6(6)	0.97

5.3 与其他无参考评价算法结果的对比

为了证明本文算法的优越性,首先在第一部分数据集的基础上与其他视频稳定评价算法进行了比较.本文选择两种经典的无参考视频稳定评价算法:帧间变换保真度(ITF)^[9]和低/高频比例(LHR)^[10]算法.其中 ITF 算法主要通过帧间峰值信噪比(PSNR)衡量两帧之间的相似性,但这种算法在具有剧烈抖动的视频类型上表现很差.LHR 算法计算运动的低频部分所占比例,本文实验中通过计算最低 6 个频率分段占全部频率分段的能量百分比

来衡量视频稳定程度.但高/低频率分隔的阈值对于不同类型的视频应该是不同的,阈值设定对算法影响很大,因此 LHR 算法无法适应不同抖动运动的视频.相比于这两种算法,本文算法直接测量运动路径的平滑程度,更加鲁棒.表 2 和表 3 分别展示了 LHR 算法和 ITF 算法与人眼主观评价的结果对比,表 4 展示了本文算法与 LHR 算法和 ITF 算法相关系数的比较,可以看出本文算法与人眼主观评测结果具有更高的一致性.

表 2 LHR 算法得到的不同稳定算法的排序结果(括号中的排序是人眼主观评价结果)

视频类别	STO	AE	YouTube	Deshaker	BP	原始抖动视频	相关系数
crowd	4(1)	3(4)	1(3)	5(5)	2(2)	6(6)	0.60
driving	3(1)	1(3)	4(4)	5(5)	2(2)	6(6)	0.77
parallax	4(2)	1(4)	3(3)	5(5)	2(1)	6(6)	0.60
rotation	6(6)	3(3)	2(1)	4(4)	1(2)	5(5)	0.94
regular	1(1)	4(4)	3(3)	5(5)	2(2)	6(6)	1.00
running	4(3)	1(4)	2(2)	5(5)	3(1)	6(6)	0.60
zooming	1(1)	2(4)	4(3)	5(5)	3(2)	6(6)	0.83
平均	4(2)	1(4)	3(3)	5(5)	1(1)	6(6)	0.76

表 3 ITF 算法得到的不同稳定算法的排序结果(括号中的排序是人眼主观评价结果)

视频类别	STO	AE	YouTube	Deshaker	BP	原始抖动视频	相关系数
crowd	3(1)	2(4)	5(3)	4(5)	1(2)	6(6)	0.60
driving	2(1)	1(3)	4(4)	5(5)	3(2)	6(6)	0.83
parallax	4(2)	2(4)	5(3)	6(5)	1(1)	3(6)	0.37
rotation	6(6)	5(3)	3(1)	1(4)	2(2)	4(5)	0.49
regular	3(1)	1(4)	4(3)	5(5)	2(2)	6(6)	0.60
running	3(3)	1(4)	4(2)	5(5)	2(1)	6(6)	0.60
zooming	4(1)	2(4)	3(3)	5(5)	1(2)	6(6)	0.60
平均	3(2)	2(4)	4(3)	5(5)	1(1)	6(6)	0.58

表 4 不同视频稳定质量评价算法对应不同类别视频的平均斯皮尔曼相关系数

视频类别	ITF	LHR	本文算法
crowd	0.60	0.60	0.94
driving	0.83	0.77	0.94
parallax	0.37	0.60	1.00
rotation	0.49	0.94	0.94
regular	0.60	1.00	1.00
running	0.60	0.60	1.00
zooming	0.60	0.83	1.00
平均	0.58	0.76	0.97

视频稳定性是视频质量好坏的衡量标准之一,但当前视频质量评价中一般只关注视频失真情况和噪声情况,并没有考虑视频稳定程度.视频稳定程度不仅影响观看者的感官体验,对于一些视频后处理工作,例如目标检测和运动跟踪等都会有严重的影响.本文选择 Wu 等人提出基于重构的无参考视频质量评价方法进行对比实验,结果如表 5 所示.其中括号内的数值是 Wu 等人算法的评价结果,它表示视频的失真程度,数值越小失真程度越小.从表中数

据可以看出 AE 稳定器和 YouTube 稳定器的大多数结果视频的失真程度均大于原始抖动视频,这主要是由于这两种稳定器在平滑轨迹后带来的影响.但在此数据集上的原始视频抖动程度非常剧烈,因

此主要影响感官体验的是抖动程度而不是失真程度,所以这里衡量视频质量的关键因素是稳定性.而本文算法在稳定性度量方面明显优于 Wu 等人提出的算法.

表 5 与无参考视频质量评价算法结果对比(括号外是本文算法得到的排序结果,括号中的数值是 Wu 等人方法的评价结果)

视频类别	STO	AE	YouTube	Deshaker	BP	原始抖动视频
crowd	1(0.12)	3(0.26)	4(0.30)	5(0.13)	2(0.17)	6(0.25)
driving	1(0.11)	4(0.18)	3(0.19)	5(0.14)	2(0.16)	6(0.21)
parallax	2(0.11)	4(0.29)	3(0.31)	5(0.13)	1(0.18)	6(0.28)
rotation	6(0.18)	3(0.38)	2(0.41)	4(0.19)	1(0.23)	5(0.36)
regular	1(0.12)	4(0.23)	3(0.25)	5(0.13)	2(0.14)	6(0.22)
running	3(0.12)	4(0.32)	2(0.35)	5(0.15)	1(0.18)	6(0.31)
zooming	1(0.13)	4(0.27)	3(0.27)	5(0.18)	2(0.16)	6(0.26)
平均	2(0.13)	4(0.28)	3(0.30)	5(0.15)	1(0.17)	6(0.27)

5.4 与全参考视频稳定评价算法结果的对比

为了进一步验证本文算法的效果,在原有数据集的基础上增加了一类视频,这类视频是真实稳定的视频,里面包含好莱坞电影场景中固定镜头拍摄的视频、扫动拍摄的视频和走动拍摄的视频.

首先通过合成抖动的方式获得抖动视频,合成

的方式主要是在 x 和 y 方向增加随机位移以及在抖动中包含尺度和旋转的变换.将增加的抖动模型按照抖动程度分为 1~4,其中 1 最轻微,4 最严重.本文算法对这一组视频的评价结果如表 6 所示.可以看出随着抖动越来越严重,本文算法计算的视频平均曲率也随着增大,表示视频越不稳定,这与实际情况相符.

表 6 本文算法对不同抖动程度视频的评价结果(抖动模型按照抖动程度分为 1~4,其中 1 最轻微,4 最严重)

视频类别	视频名称	稳定视频	抖动 1	抖动 2	抖动 3	抖动 4
静态相机	Video1	1.12	1.31	1.44	1.54	1.65
	Video2	1.15	1.36	1.45	1.58	1.64
	Video3	1.10	1.35	1.48	1.57	1.68
	Video4	1.11	1.34	1.47	1.54	1.66
	Video5	1.12	1.33	1.47	1.55	1.64
扫动相机	Video6	1.17	1.38	1.48	1.59	1.69
	Video7	1.15	1.36	1.46	1.61	1.66
	Video8	1.18	1.39	1.49	1.57	1.66
	Video9	1.17	1.40	1.51	1.60	1.68
	Video10	1.18	1.35	1.47	1.58	1.65
走动相机	Video11	1.21	1.41	1.52	1.63	1.68
	Video12	1.19	1.38	1.51	1.61	1.71
	Video13	1.23	1.39	1.51	1.58	1.73
	Video14	1.20	1.42	1.49	1.59	1.72
	Video15	1.21	1.39	1.53	1.58	1.71
平均		1.17	1.37	1.49	1.58	1.68

然后我们使用 5 种视频稳定算法对合成的抖动视频进行处理进而获得稳像后的视频,用于与全参考视频稳定质量评价算法作对比.本文使用典型的全参考视频稳定质量评价算法^[6],即稳像后的路径与原始稳定路径的距离来衡量视频稳定程度,记为 d_{path} ,它是通过计算两个视频每帧的路径距离的差异绝对值,最后求平均得到. d_{path} 值越小,表明稳像后的视频越接近真实的稳定视频,如下式:

$$d_{path} = \frac{1}{N} \sum_{t=1}^N |Path_s(t) - Path_o(t)| \quad (13)$$

其中, N 表示视频帧数, $Path_s$ 表示稳像后视频路径, $Path_o$ 表示真实稳定视频的路径.

本文算法对包含真实稳定视频的数据集的评价结果与全参考方法的评价结果如表 7 所示,表中括号外的数值是本文计算得到的平均曲率结果,括号内的数值是全参考方法计算得到的平均路径差异值,最后一列是这两种评价结果的相关系数.从表 7 可以看出,本文算法与全参考方法的结果相较于人眼主观评价结果一致性有所降低,但依然具有较高的一致性.并且可以看出全参考方法计算 YouTube

算法稳像结果与真实稳定视频的平均路径距离有时甚至大于抖动视频与真实稳定视频的平均路径距

离,而这时本文的无参考方法仍能获得正确的评价结果.

表 7 本文算法得到的平均曲率结果(括号中的数值是全参考方法获得的稳像视频与真实稳定视频的平均路径距离)

视频类别	视频名称	稳定视频	STO	AE	YouTube	Deshaker	BP	原始抖动视频	相关系数
静态相机	Video1	1.12(0)	1.31(18.28)	1.33(27.25)	1.29(21.98)	1.54(29.66)	1.27(24.56)	1.59(44.72)	0.86
	Video2	1.15(0)	1.23(41.82)	1.35(59.83)	1.25(36.57)	1.58(69.15)	1.28(32.79)	1.60(89.54)	0.86
	Video3	1.10(0)	1.35(28.23)	1.33(27.88)	1.32(34.32)	1.39(29.84)	1.25(26.26)	1.47(31.73)	0.64
	Video4	1.11(0)	1.36(19.76)	1.40(22.10)	1.41(31.27)	1.45(24.86)	1.33(18.07)	1.51(29.53)	0.89
	Video5	1.12(0)	1.26(12.87)	1.34(16.73)	1.29(19.34)	1.37(19.80)	1.24(13.35)	1.43(22.92)	0.79
抖动相机	Video6	1.17(0)	1.22(17.69)	1.28(23.71)	1.26(27.06)	1.36(26.53)	1.24(18.97)	1.50(28.95)	0.89
	Video7	1.15(0)	1.28(33.75)	1.35(36.74)	1.33(37.56)	1.44(39.86)	1.25(31.94)	1.53(45.50)	0.96
	Video8	1.18(0)	1.31(23.17)	1.33(25.80)	1.34(38.12)	1.38(26.72)	1.27(22.53)	1.48(31.86)	0.89
	Video9	1.17(0)	1.23(45.24)	1.48(54.68)	1.29(327.8)	1.53(59.32)	1.26(49.04)	1.65(76.39)	0.79
	Video10	1.18(0)	1.27(43.57)	1.39(54.59)	1.34(46.98)	1.46(55.20)	1.26(41.05)	1.57(58.29)	1.00
走动相机	Video11	1.21(0)	1.34(39.87)	1.41(47.73)	1.37(54.28)	1.57(58.85)	1.28(38.76)	1.62(68.09)	0.96
	Video12	1.19(0)	1.38(38.86)	1.43(49.82)	1.36(45.16)	1.53(53.18)	1.33(32.92)	1.55(54.72)	0.96
	Video13	1.23(0)	1.26(72.53)	1.43(85.78)	1.39(143.6)	1.59(88.31)	1.28(76.83)	1.67(97.34)	0.79
	Video14	1.20(0)	1.26(29.12)	1.34(36.49)	1.32(39.94)	1.37(38.75)	1.29(32.37)	1.54(41.72)	0.89
	Video15	1.21(0)	1.33(55.52)	1.46(59.26)	1.40(63.56)	1.48(73.91)	1.36(53.71)	1.64(84.57)	0.93
平均									0.87

6 结 语

我们提出一种新的基于运动路径全曲率的无参考视频稳定质量评价算法.从运动感知的观点来看,平滑的运动一般具有较好的曲率特性,这让本文算法的评价结果与人眼主观评价结果具有很高的一致性,实验也证明了我们的无参考评价算法的有效性.

运动稳定性事实上是一种视觉感知驱动的现象,而运动路径全曲率正是运动感知在几何上的体现.但全曲率并不是影响视频稳定性的唯一感知属性,因此在下一步的工作中,我们将加入更多的感知属性,进一步提高视频稳定质量评价与人眼感知的一致性.

参 考 文 献

- [1] Grundmann M, Kwatra V, Essa I. Auto-directed video stabilization with robust L1 optimal camera paths//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA, 2011: 225-232
- [2] Liu F, Gleicher M, Jin H, et al. Content-preserving warps for 3D video stabilization. ACM Transactions on Graphics, 2009, 28(3): 44
- [3] Liu F, Gleicher M, Wang J, et al. Subspace video stabilization. ACM Transactions on Graphics, 2011, 30(1): 4
- [4] Zhang L, Xu Q K, Huang H. A global approach to fast video stabilization. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(2): 225-235

- [5] Tanakian M J, Rezaei M, Mohanna F. Camera motion modeling for video stabilization performance assessment//Proceedings of the IEEE Machine Vision and Image Processing. Tehran, Iran, 2012: 1-4
- [6] Qu H, Song L, Xue G. Shaking video synthesis for video stabilization performance assessment//Proceedings of the IEEE Visual Communications and Image Processing. Kuching, Malaysia, 2013: 1-6
- [7] Wu Z, Hu H. Reconstruction-based no-reference video quality assessment//Proceedings of the IEEE Region 10 Conference. Singapore, 2017: 3075-3078
- [8] Lei Z, Chen X Q, Kong X Y, et al. Geodesic video stabilization in transformation space. IEEE Transactions on Image Processing, 2017, 26(5): 2219-2229
- [9] Battiato S, Gallo G, Puglisi G, et al. SIFT features tracking for video stabilization//Proceedings of the IEEE International Conference on Image Analysis and Processing. Modena, Italy, 2007: 825-830
- [10] Liu S, Yuan L, Tan P, et al. Bundled camera paths for video stabilization. ACM Transactions on Graphics, 2013, 32(4): 1-10
- [11] Panetta K, Gao C, Aгаian S. No-reference color image contrast and quality measures. IEEE Transactions on Consumer Electronics, 2013, 59(3): 643-651
- [12] Offiah M C, Amin N, Gross T, et al. An approach towards a full-reference-based benchmarking for quality-optimized endoscopic video stabilization systems//Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing. New York, USA, 2012: 1-8
- [13] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features. Computer Vision and Image Understanding, 2006, 110(3): 404-417

- [14] Zhang Kun, Xu Ting-Fa, Wang Ping, et al. Real-time full-frame digital image stabilization system by SURF. *Optics and Precision Engineering*, 2011, 19(8): 1964-1972(in Chinese) (张坤, 许廷发, 王平等. 高精度实时全帧频 SURF 电子稳像方法. *光学精密工程*, 2011, 19(8): 1964-1972)
- [15] Hartley R, Zisserman A. *Multiple View Geometry in Computer Vision*. New York, USA: Cambridge University Press, 2003
- [16] Helgason S. *Differential Geometry, Lie Groups and Symmetric Spaces*. New York, USA: Academic Press, 1978
- [17] Peli E, García-Pérez M A. Motion perception during involuntary eye vibration. *Experimental Brain Research*, 2003, 149(4): 431-438
- [18] López M C, Mateos V F, Masqué J M. Total curvature of curves in riemannian manifolds. *Differential Geometry and Its Applications*, 2010, 28(2): 140-147
- [19] Zacur E, Bossa M N, Olmos S. Left-invariant riemannian geodesics on spatial transformation groups. *SIAM Journal on Imaging Sciences*, 2014, 7(3): 1503-1557
- [20] Wang Y S, Liu F, Hsu P S, et al. Spatially and temporally optimized video stabilization. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(8): 1354-1361
- [21] Marden J I, Kendall M, Gibbons J D. Rank correlation methods. *Journal of the American Statistical Association*, 1992, 87(417): 249



ZHENG Qing-Zhuo, born in 1992, M. S. candidate. His research interests include image and video processing.

ZHANG Lei, born in 1981, Ph. D., associate professor. His research interests include computer graphics, image and video processing.

HUANG Hua, born in 1975, Ph. D., professor. His research interests include image and video processing.

Background

In daily life, due to the poor shooting environment and the lack of professional equipment, often cause serious jitter in the video obtained, affecting the video quality and user experience. As a kind of technology that can remove the video jitter and enhance the visual quality, video stabilization has been widely and deeply studied, and a large number of video stabilization algorithms have been proposed in the last ten years. However, how to evaluate the performance of video stabilization algorithms effectively is a problem that has not been studied deeply. Due to the lack of a comprehensive and fair evaluation mechanism, it is impossible to evaluate the performance of video stabilization algorithm, which makes it difficult for users to choose video stabilization algorithms.

There are two kinds of evaluation methods: non-reference and full-reference in the field of video stabilization quality assessment. Full-reference method requires a real stable video as a benchmark, but this benchmark video is difficult to

obtain. Non-reference video quality assessment method obtains the jitter characteristics of the video by using the statistical model, which can be used to evaluate video stabilization algorithm. This method is flexible and economical, so this paper chooses the non-reference method to evaluate the video stabilization algorithm objectively. Motion stability is essentially a kind of perceptual attribute, which is difficult to be modeled mathematically. But from the point of view of motion trajectory, the trajectory of stabilization video is generally smooth, and the motion trajectory of the shaky video generally has more peaks and troughs. In mathematics, curvature is used to measure the degree of curvature of a curve. So, this paper studies the no-reference video stabilization quality assessment directly from the degree of motion smoothness, proposes a no-reference video stabilization evaluation method based on the curvature of the motion path.