

文章编号:1002-8692(2011)07-0015-03

# 视频稳像技术综述

陈启立, 宋 利, 余松煜

(上海交通大学 电子工程系 图像通信与信息处理研究所;上海市数字媒体处理与传输重点实验室,上海 200240)

**【摘 要】** 分析了视频稳像的各类算法,包括基于运动的经典算法框架下的各种研究和无运动的新稳像算法。指出了各类算法的优缺点和适用性,并展望了今后的研究方向。

**【关键词】** 视频稳像;运动估计;运动补偿;特征点

**【中图分类号】** TN911.73;TP391

**【文献标识码】** A

## An Overview of Video Stabilization

CHEN Qili, SONG Li, YU Songyu

(Institute of Image Communication and Information Processing, Department of Electronic Engineering, Shanghai Jiao Tong University;  
Shanghai Key Laboratory of Digital Media Processing and Transmissions, Shanghai 200240, China)

**【Abstract】** Different video stabilization algorithms are analyzed in this paper, including motion-based algorithms under the classical stabilization framework and the novel non-motion algorithms. The applicability and features of the algorithms are also presented. At the end, the existing problems and future prospects are pointed out.

**【Key words】** video stabilization; motion estimation; motion compensation; feature points

## 0 引言

如今,图像视频无处不在,人们几乎天天都会与视频打交道,生活离不开电视电影,社会的治安离不开监控。而随着媒体技术的发展,摄像设备越来越普及,越来越多的普通家庭及个人有条件自行拍摄、处理各类视频,而不仅仅局限于观看。在这一形势下,视频质量备受关注,针对手持摄像机及车载摄像设备,视频抖动问题显得尤为突出。在飞机、坦克等军用场合,同样十分有必要对视频进行稳定化处理。视频稳定(简称稳像)就是在这些需求下诞生的。所谓稳像,是指利用相关的算法,对视频设备采集的原始视频序列进行处理,去除其中的抖动。视频稳像的目的,一方面是为了让人眼观感舒适,有利于人工观测、判别等,另一方面也作为诸多其他后续处理的预处理阶段,如检测、跟踪和压缩。笔者将综述视频稳像的各类经典算法和近年来取得突破成果的算法。

## 1 算法分类与经典算法

### 1.1 稳像算法分类

稳像按作用机制分为光学、机械和电子稳像。光学稳像和机械稳像牵涉到物理设备,主要为摄像机产商和专业摄影者使用,也有文章对这类稳像算法作分析和研究<sup>[1]</sup>。光学稳像的基本途径是通过光学设备调整光路从

而补偿摄像平台带来的抖动;机械稳像是利用各种机械设备来设法保证摄像平台的稳定性,在影像拍摄上有较多应用。以上两类稳像对普通消费者而言,代价过大,而且机械稳像设备不便携带;另一方面,光学稳像和机械稳像所能达到的效果仍不能令人满意,尤其是当抖动较为剧烈时,从而使得电子稳像变得十分有必要。

### 1.2 基于运动的稳像算法

经典的电子稳像包含3个步骤:全局运动估计、运动补偿和图像生成。各种算法都遵循这个框架,在每一个步骤里,又有各自不同的实现。全局运动一般指的是视频中背景的运动。在做全局运动估计时,一个难点是如何获得准确的运动信息,排除局部运动等各种干扰的影响。主要途径有微分方法和特征点对应法。微分方法如光流法<sup>[2]</sup>:首先计算光流,再根据光流来求帧间几何关系,以此来表示全局运动。基于特征点的方法是通过特征点的对应关系来描述运动,同样利用图像帧间几何关系等来表示运动,常用的特征有边缘、SIFT<sup>[3]</sup>,也有算法基于局部运动矢量进行块匹配<sup>[4]</sup>。它们的思想都是利用局部量,约定目标函数,求取表示全局运动的参数(如变换矩阵),在精度要求不是很高的场合下,仿射变换矩阵经常被选用。

运动补偿是指对全局运动进行修正,使主观运动与抖动分离,是视频稳像的实质所在。如图1所示,其目标是生成平滑的运动。运动补偿分为2类:基于参数滤波和基于轨迹平滑。基于参数滤波是指把描述运动的参数

基金项目:国家自然科学基金项目(60702044;60625103)

看成需要的主观运动和加性抖动的叠加,采用一定的滤波方法使得加性噪声运动被抑制,如采用概率模型的Kalman滤波<sup>[6]</sup>。文献[6]构造状态空间模型来描述全局运动和抖动的参数,并进一步将去抖问题转化为Kalman滤波问题,以此来求取所需变换矩阵的参数。基于轨迹平滑则将摄像机的运动轨迹看成带噪声的运动轨迹<sup>[7]</sup>,采用相应的平滑方法来去除高频噪声。文献[7]采用了基于贝塞尔(Bezier)曲线的轨迹滤波方法,它最大的特点是平滑轨迹的同时,能确保输出的视频帧在拼接合成的扩展图像范围内,从而使得视频质量上有较好保证。运动补偿还有一些特别的处理方法,如把去抖动问题转化为控制问题<sup>[8]</sup>,文献[8]利用修正的比例积分控制器来求取运动补偿向量。

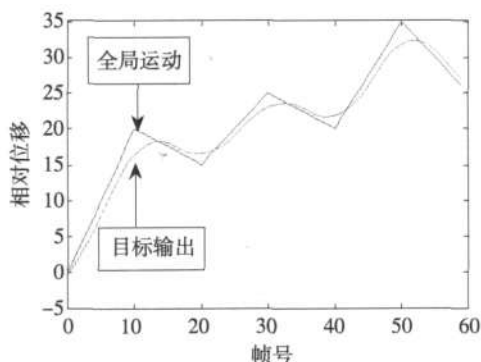


图1 运动补偿示意图

图像生成是稳定算法的后续阶段,其目标是根据补偿后的运动生成稳定视频输出流,涉及到图像拼接、去模糊、融合等技术。相对简单的实现是输出与输入帧一对一变换<sup>[6]</sup>,如图2所示。在这种模式下,输出视频时会出现空白区域。许多学者也研究了如何保持图像精度(全帧输出)。通常的方法是通过相邻帧的拼接来获得全帧输出,其缺点是可能会使得拼接区域过渡不连续、不自然。文献[9]提出了一种通过图像修补获得全帧输出的方法。其思想是:首先求取局部运动矢量,将局部运动矢量延伸到空白区域,再利用所获得的信息来引导变换矩阵的生成。其他方法还有通过像素填充<sup>[10]</sup>等。在图像生成时另一个思路是把这个过程看成一个渲染生成的过程<sup>[11]</sup>,其思想是综合图像信息,在此基础上根据特定的约束和一定的准则生成图像,而不是局限于图像的一一对应关系。如图3所示,该方法不是一对一的图像变换,而是基于一个子序列。将子序列送渲染系统,经过插值、融合等操作,生成不同视点的输出序列。这种方法的优点是生成的视频序列信息更全,也有更好的连贯性与视觉效果。

什么是稳定的视频,应该是稳像首先应该回答的问题。在经典的算法中,多数将其归结为运动的平滑;文

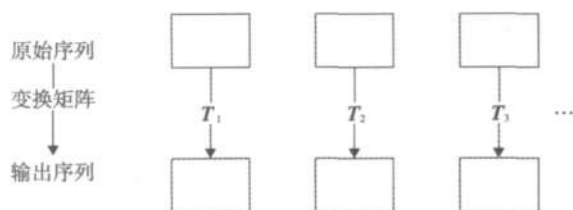


图2 常规图像生成示意图

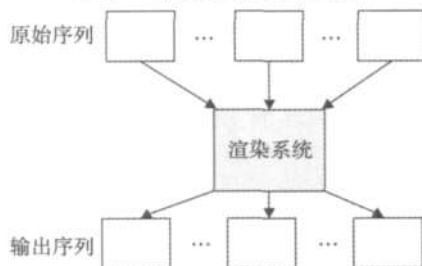


图3 基于渲染的图像生成示意图

献[12]提出了新的准则,以图像内容作为评价视频稳定的标准,以此来引导稳像。近年来,还有学者提出了无运动的稳像算法,从而稳定的准则也不再是运动的平滑,取而代之的是其他约束条件。

## 2 无运动的稳像算法

传统的视频稳定算法都是建立在第1节所述的框架下。然而在这种基于运动的稳像框架下,有着部分无法解决的问题:

1) 全局运动本身很难精确地获得。稳像算法能取得效果的首要条件是准确的全局运动估计,但所需的全局运动是未知的。许多算法都是建立在二维模型上,本身就难以准确求解三维场景的帧间关系;另外,运动是通过特征等中间桥梁关联起来的,即使帧间关系能完全准确地得出(如变换矩阵),其与全局运动并不是完全等价的。

2) 运动补偿难以很好地满足视觉效果的去抖动。做运动补偿时,与问题1)相似,对参数滤波,特征轨迹平滑,不能直接与运动平滑等价起来;另一方面,运动补偿时普遍存在误差累积问题,有学者用周期矫正策略来减少累积误差的影响<sup>[13]</sup>,但不能从根本上消除误差传播的影响。

以上原因使得传统算法难以解决剧烈抖动、动态场景、多深度变化场景、多运动物体、近距运动物体等一个或多个挑战。在这种形势下,如何改变或跳出传统算法的框架,建立新的稳像模型成了解决这些挑战的关键所在。

近年来,有学者针对以上问题,做了深入研究,并取得了十分显著的效果。它们有光流视频稳像<sup>[14]</sup>和基于稳健(robust)特征轨迹的视频稳像<sup>[15]</sup>。其中文献[14]和[15]都没有显式的运动估计,取而代之的是其他准则与处理。与传统算法相比,算法框架有了新的突破,稳像效果也有了显著提升。

## 2.1 光流稳像

光流稳像中的光流指的是输入视频来自光流摄像机,即由几何关系已知的摄像机阵列输入多路同时不同视点的视频流。在这种条件输入下,图像三维坐标能够准确求取,具体方法是通过多视输入获得的深度信息<sup>[16]</sup>求出三维坐标。稳像的基本思路是直接由得出的三维位置在特定的约束条件下生成稳定视频序列。稳定序列的与参考序列的关系未知,在约束条件下求最优解得出。完整过程是:1) 计算每一帧的深度信息;2) 计算每一帧的光流;3) 检测 canny 边缘,利用光流进行匹配;4) 求取约束最优解(即目标序列与参考序列的关系);5) 图像生成。结合以上分析,此算法的优点是:首先,它能获得准确的三维空间坐标信息;第二,没有运动估计,直接由输入序列在约束准则下得到稳定输出序列,从而跳过了运动估计、运动补偿等步骤,有效解决了传统算法的问题,所以有十分好的稳定效果。在剧烈抖动、动态场景、多深度变化场景、多运动物体、近距运动物体的场景中,均有比较好的效果。主要的缺点是计算复杂度较高,另外阵列摄像机对普通消费者来说仍是奢侈品。

## 2.2 基于稳健特征轨迹

基于稳健特征轨迹的算法中同样没有运动估计和运动补偿,它用特征轨迹作为无参的运动替代量,并将稳像问题融合到优化问题中。基本步骤为:1) 稳健特征轨迹求取(分为特征点添加、连接、轨迹延伸和轨迹删减4个子过程);2) 优化稳像。优化的目标函数是包含表征轨迹光滑度和图像失真度的2个能量函数,通过1个可调参数来权衡两者间的权重。在这一目标函数的约束下,使得最终获得的输出序列在稳定化的同时有较高的视频质量。其缺点是,视频序列中必须有足够的可用特征点来保证算法的有效性。

## 2.3 无运动稳像的优势

无运动稳像算法代表着视频稳像领域的一个新高度,它们都取得了很好的实践效果,能够解决剧烈抖动、动态场景、多深度变化场景、多运动物体、近距运动物体诸多挑战中的数个挑战。相比基于运动的经典稳像算法,其优势有:1) 跳出了传统的视频稳像框架。都没有运动估计和运动补偿,而运动正是经典算法取得突破的瓶颈所在。2) 新的稳定概念。什么是稳定的视频,应该是稳像首先应该回答的问题。在经典的算法中,多数将自归结为运动的平滑。基于稳健特征的稳像用特征轨迹的光滑性来表征稳定序列,以此来引导稳像,这一准则和人眼的观感有着更直接和紧密的联系。3) 利用了更多的信息。光流稳像属于三维稳像,人眼所获得的视

觉信息本身就是三维的,三维稳像提供了更多、更准确的位置信息,有利于稳像获得更佳效果。

## 3 总结与展望

得益于技术的发展,人们能越来越方便地拍摄和取得视频。从而使得视频质量更受关注,视频稳像能够帮助业余人士获得赏心悦目的视频。而视频稳像是一个反问题,本身并没有唯一解。该问题已知输入,经过一系列步骤,规定目标和相关约束,进行优化求解。虽然稳像技术取得了很好的发展,尤其是各类新兴算法的出现。但依然存在一些问题有待解决:

1) 难以有算法适用于各种场景。有些算法要求有足够多的特征点,有些算法只能有效用于静态场景。尤其是对于剧烈抖动、动态场景、多深度变化场景、多运动物体、近距运动物体等挑战性场景,鲜有算法能同时解决。

2) 稳像效果和稳像效率上难以兼得。光流稳像虽然能处理上述的多数挑战,但是对于普通分辨率(480×360)的视频,每帧需要十几分钟的处理时间。相当于1 min的序列需要数天的处理时间,这对用户来说是难以容忍的。而诸多高效算法难以实现对各种场景均有令人信服的效果。

3) 在稳定视频的评估上,没有有效的标准,虽然也有文章对此进行专门研究<sup>[17]</sup>。常用的有峰值信噪比(PSNR)、偏移范围等评价方法,但这些方法都难以很好地与视觉效果直接相关联。所以,到目前为止,最有效的方法依然是直接由人眼来评定。而这种主观的方法十分消耗精力,且有很大的不确定性。另外,尚未建立公认的公共测试集。

视频稳像的发展方向也就是如何解决上述问题。首先需要建立合理丰富的测试集,供广大研究人员共同使用评估,同时要探索有效的评估方法。然后是探索新的方法,力求做到效果和效率上的同时提升,最终的目标是实时的稳健稳像。

## 参考文献:

- [1] SACHS D, NASIRI S, GOEHL D. Image stabilization technology overview[M]. Santa Clara, CA: InvenSense Inc., 2007:18.
- [2] CHANG H C, LAI S H, LU K R. A robust real-time video stabilization algorithm[J]. Journal of Visual Communication and Image Representation, 2006, 17(3): 659-673.
- [3] BATTIATO S, GALLO G, PUGLISI G, et al. Sift features tracking for video stabilization[EB/OL]. [2010-04-06]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.144.3804&rep=rep1&type=pdf>.
- [4] BATTIATO S, PUGLISI G, BRUNA A R. A robust video stabilization system by adaptive motion vectors filtering[EB/OL]. [2010-04-06].

(下转第22页)



总之,CSURF方法是一种稳健性很高的图像配准算法,为彩色图像的配准问题提供一种新的解决方法。

#### 参考文献:

- [1] HAIKEL A, MOHAMD K. Image registration using virtual circles and edge direction[C]//Proc. IEEE 16th International Conference on Pattern Recognition. Quebec, Canada: IEEE Press, 2002: 969-972.
- [2] ZITOVA B, FLUSSER J. Image registration methods: a survey[J]. Image and Vision Computing, 2003, 21(11): 977-1000.
- [3] BROWN L G. A survey of image registration techniques[J]. ACM Computing Surveys, 1992, 24(4): 325-376.
- [4] 倪国强, 刘琼. 多源图像配准技术分析与展望[J]. 光电工程, 2004, 31(9): 1-6.
- [5] LOWE D. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [6] BAY H, TUYTELAARS T, GOOL L. SRUF: speeded up robust features [C]//Proc. the 9th European Conference Computer Vision. [S.l.]: IEEE Press, 2006: 404-417.
- [7] 郝明非, 张建秋, 胡波. 一种超复数鲁棒相关图像配准算法[J]. 复旦学报: 自然科学版, 2007, 46(1): 91-95.
- [8] 冯巍, 胡波. 基于超复数相位相关的彩色图像配准方法[J]. 系统工程

与电子技术, 2010, 32(1): 183-187.

- [9] ALAA A, ALY F. CSIFT: A SIFT descriptor with color invariant characteristics[C]//Proc. the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Press, 2006: 1978-1983.
- [10] 张锐娟, 张建奇, 杨翠, 等. 基于CSIFT的彩色图像配准技术研究[J]. 光学学报, 2008, 28(11): 2097-2103.
- [11] GEUSEBROEK J, BOOMGAARD R. Color invariance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(12): 133.
- [12] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Communication of ACM, 1981, 24(6): 381-395.



#### 作者简介:

刘学(1983-), 硕士生, 主研图像配准算法;  
姚洪利(1972-), 高工, 主要研究方向为视频图像处理、图像处理等;  
金世龙(1965-), 博士生导师, 主要研究方向为光学陀螺、光学加工、数字图像处理等。

责任编辑: 许盈

收稿日期: 2010-07-07

(上接第17页)

[http://iplab.dmi.unict.it/download/Elenco%20Pubblicazioni%20\(PDF\)/International%20Conferences/IC63%20IEEE%20ICME08%20Block\\_Based\\_Video\\_Stabilization.pdf](http://iplab.dmi.unict.it/download/Elenco%20Pubblicazioni%20(PDF)/International%20Conferences/IC63%20IEEE%20ICME08%20Block_Based_Video_Stabilization.pdf).

- [5] ZHANG Y, ZHANG W, ZHAO X, et al. Study on electronic image stabilization system based on MEMS Gyro[EB/OL]. [2010-04-06]. <http://www.computer.org/portal/web/csd/doi/10.1109/ICECT.2009.27>.
- [6] LITVIN A, KONRAD J, KARL W C. Probabilistic video stabilization using Kalman filtering and mosaicing[EB/OL]. [2010-04-06]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.1398&rep=rep1&type=pdf>.
- [7] YAN W, KANKANHALLI M S. Detection and removal of lighting & shaking artifacts in home videos[EB/OL]. [2010-04-06]. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.65.7800>.
- [8] LIN C T, HONG C T, YANG C T. Real-time digital image stabilization system using modified proportional integrated controller [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(3): 427-431.
- [9] MATSUSHITA Y, OFEK E, TANG X, et al. Full-frame video stabilization[EB/OL]. [2010-04-06]. [http://research.microsoft.com/en-us/people/yasumat/fullframe\\_cvpr05.pdf](http://research.microsoft.com/en-us/people/yasumat/fullframe_cvpr05.pdf).
- [10] CHUNG C Y, CHEN H H. Feature-based full-frame image stabilization[C]//Proc. ISM 2007. Taichung: IEEE Press, 2007: 100-106.
- [11] BUEHLER C, BOSSE M, MCMILLAN L. Non-metric image-based rendering for video stabilization[EB/OL]. [2010-04-06]. <http://www.computer.org/portal/web/csd/doi/10.1109/CVPR.2001.991019>.

991019.

- [12] LIU F, GLEICHER M, JIN H, et al. Content-preserving warps for 3D video stabilization[EB/OL]. [2010-04-06]. [http://www.cs.wisc.edu/graphics/Papers/Gleicher/fliu/siggraph09\\_preprint.pdf](http://www.cs.wisc.edu/graphics/Papers/Gleicher/fliu/siggraph09_preprint.pdf).
- [13] LUO Q, KHOSHGOFTAAR T M. An empirical study on estimating motions in video stabilization[C]//Proc. IRI 2007. Las Vegas, IL: IEEE 2007: 360-366.
- [14] SMITH B M, ZHANG L, JIN H, et al. Light field video stabilization [EB/OL]. [2010-04-06]. <http://pages.cs.wisc.edu/~lizhang/projects/lfstable/SmithICCV09.pdf>.
- [15] LEE K, CHUANG Y, CHEN B, et al. Video stabilization using robust feature trajectories[EB/OL]. [2010-04-06]. <http://www.csie.ntu.edu.tw/~cyy/publications/papers/Lee2009VSR.pdf>.
- [16] SMITH B M, ZHANG L, JIN H. Stereo matching with nonparametric smoothness priors in feature space[EB/OL]. [2010-04-06]. [http://pages.cs.wisc.edu/~lizhang/projects/mvstereo/cvpr2009/SmithCVPR09\\_presentation.pdf](http://pages.cs.wisc.edu/~lizhang/projects/mvstereo/cvpr2009/SmithCVPR09_presentation.pdf).
- [17] MORIMOTO C, CHELLAPPA R. Evaluation of image stabilization algorithms[EB/OL]. [2010-04-06]. [http://www.netsoc.tcd.ie/~fastnet/cd\\_paper/ICASSP/ICASSP\\_1998/pdf/author/ic981849.pdf](http://www.netsoc.tcd.ie/~fastnet/cd_paper/ICASSP/ICASSP_1998/pdf/author/ic981849.pdf).



#### 作者简介:

陈启立(1986-), 硕士, 主研视频稳像;  
宋利, 副教授, 研究领域为视频编码、视觉计算;  
余松煜, 教授, 博士生导师, 研究方向为图像通信、数字电视、图像处理 and 模式识别。

责任编辑: 哈宏疆

收稿日期: 2010-11-01