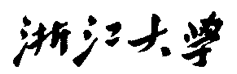
他们硕士研究读书报告





题 目 视频时空一致性

作者姓名 徐 辉

作者学号 NB15156

指导教师 李启雷

学科专业 移动互联网与游戏开发

所在学院 软件学院

提交日期 二○一五年十二月

The Consistency of Video Time

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Mobile Internet and game development

Advisor: Li Qilei

By

Xu Hui

Zhejiang University， P.R. China

2015

**摘要**

扩展图像处理技术对视频来说是一个不平凡的任务，在应用处理视频每一帧时常常导致时间不一致性，并且要明确编码时间的一致性，就需要改变算法。我们以时间一致性来描述一个更通用的方法。我们提出了一种梯度域技术，它是盲目的特定图像处理算法。我们技术主要处理那些患有闪烁的和产生在时间上一致的视频序列的帧。我们的解决方案核心是从原始未经处理的视频推断时间规律性，

并把它作为一个时间一致性导向稳定处理序列算法。我们正式展示和体现我们技术的频率特性，在实际应用中，它的主要能力是稳定的广泛流行的图像处理技术，包括增强色彩和色调，固有图像的风格化，和深化估计能力。

**关键字：**视频处理；时间上的一致性。

**Abstract**

Extending image processing techniques to videos is a non-trivial task; applying processing independently to each video frame often leads to temporal inconsistencies, and explicitly encoding temporal consistency requires algorithmic changes. We describe a more general approach to temporal consistency. We propose a gradient domain technique that is blind to the particular image processing algorithm. Our technique takes a series of processed frames that suffers from flickering and generates a temporally-consistent video sequence. The core of our solution is to infer the temporal regularity from the original unprocessed video, and use it as a temporal consistency guide to stabilize the processed sequence. We formally characterize the frequency properties of our technique, and demonstrate, in practice, its ability to stabilize a wide range of popular image processing techniques including enhancement and stylization of color and tone, intrinsic images, and depth estimation.

**Keywords:** Video processing; temporal consistency.

**1. 简介**

随着加工的进步，有效滤池进行恢复和增强，创造性的编辑和分析，现在常态的静态图像。视频，而另一方面，不享受同样丰富多彩的工具箱。人们自然的把视频序列独立的沙化过程为独立的一帧帧每个带有一个滤波器用于设计的静态图片。这可能工作在一些简单的情况下，像高通和低通滤波，但在许多其它更复杂的情况下，这会产生从闪烁遭受难看的结果。发生这种情况可能由于各种原因，例如，一个复杂的优化技术可能落入不同局部极小根据在框架上，或者一个滤波器可取决于统计，像平均颜色，是不稳定的整个视频序列。

一个解决这个问题的方法是明确地考虑时间一致性。多个视频处理算法已被沿着这条线的发展，如颜色分级[Bonneel等。2013]，动态范围压缩[艾登等。2014年]，内在分解[Ye等。2014年; Bonneel等。 2014年;kong等。2014年]，和色调稳定[Farbman和Lischinski2011]。尽管有效，但这些技术是特定网络到每一个任务，不要推广到其他问题。Paris[2008]和Lang等人[2012]提出了更通用的方法，以静止图像运营商扩展到视频。然而，这些继续承担一个特定的滤波器的配方，这限制了它们的应用。对于这一套，Lang等人以外的运营商则诉诸时间低通滤波。正如他们看到的，这减少了闪烁，但并没有完全将其删除。

他们的目标是一个更一般的方法来扩展图像滤池到视频，并提出一种算法是不可知的滤波器的内部设计，也就是说，他们把图像滤池为黑色框，需要输入帧，并生成处理的帧。在这个意义上，他们的方法，以时间一致性是盲目的图像滤波器被应用。他们有两个要求：

1. 该原始视频信息，并将该光学或特征溢流是可恢复的，使得它可以用作一个时间一致性导向;
2. 该滤波器不产生新的内容不相关的输入，例如，绘画渲染滤池的程序上产生笔触的质感和补画了合成新的内容都是他们的范围之内的技术。

这就是说，他们的方法，涵盖了各种滤池，例如去雾，自动摄影增强，颜色分级以及内在的分解。他们制定他们的梯度域算法，并提出了一个能量函数，无异于空间筛选泊松方程与时间的限制，他们可以有效地解决。他们正式的频率内容方面的特点他们的方法的性能。他们提供了实验证明，它可以处理各种效果，它可以直向前延伸国家的最先进的图像滤池到视频。贡献：1.由于视频帧图像滤波器的应用消除闪烁技术。2. 一个梯度域形成的时间一致性问题无关应用图像过滤器。3. 隐式扩展的几个最先进的否不稳定的帧视频图像过滤器。

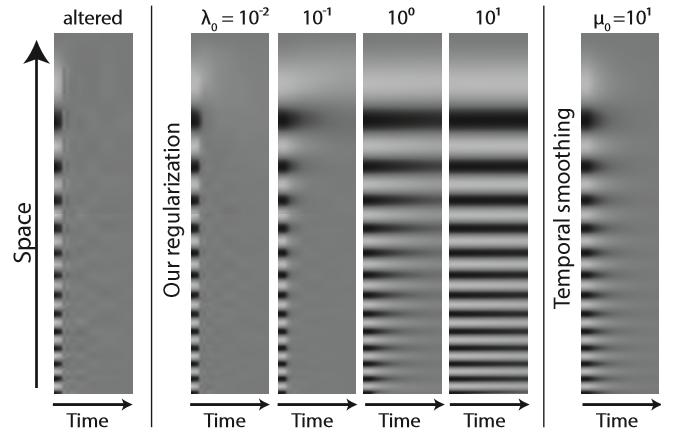
**2. 相关工作**

当应用于帧一帧，一些图像滤池，如低通滤波和边缘保持同行产生暂时稳定的结果，例如，[Winnem¨olleretal.2006; Chenetal.2007]。

然而，许多其他图像滤池不推广到视频和需要进行调整。

一种流行的解决方案是把重点放在一个给定的滤波器为一个特定的应用程序。例如，Bonneel等。 [2013]和Wang等。 [2006]通过在时间上滤波的颜色传递函数一个视频的颜色等级转移到另一个。 艾登等人[2014]进行视频色调映射在时间上分解HDR内容为基础，细节层次一致，暂时滤波基层更积极。Bonneel等。[2014年]，Ye等。[2014]，和Kong等。[2014年]产生稳定的视频内在的分解。 所有这些方法都很好地满足他们的需要，而是因为他们的方式执行时间一致性依赖于他们的目标应用程序的具体情况，他们不推广到其他应用程序。相比较而言，他们建议，适用于大量图象滤池的一种方法。更一般的方法被提出来处理滤池的全部课程。Paris[2008]延伸的高斯核到时域，并使用该结果，以适应应用，例如双边滤波和均值漂移聚类视频。Lang等。[2012]还通过利用光溢流平滑扩展到时域的概念，重温像运动估计和着色优化为基础的技术。对于不优化能源等技术，他们采取时间低通滤波。他们将在结果一节时间平滑降低了高频闪烁，但低频的不稳定性依然看。不像这两种方法，他们的方法不要求图像滤波器为具有特定的形式，也不适应其制剂。相反，他们运行不改性科幻阳离子滤波器作为一个黑盒子，打造一个集网络过滤的帧，然后对其进行处理，以去除颞文物。相反，他们运行滤波器不加修改地作为一个黑盒子，打造一个集网络过滤的帧，然后对其进行处理，以去除时间伪影。

同时这项工作，侗等。[2015]提出了一种技术来稳定由未知图像滤波器可以表示为非线性曲线应用于原始视频帧的区域处理的视频帧。相比之下，他们的技术不限于特定制剂，可以处理像征图像或深度预测违反这个假设的应用程序。平行地，技术已经发展到除去不一致输入的视频，而不是在视频一个滤波器处理。这包括从像相机自动白平衡拍摄时间问题，从不匹配的相机/照明/投影机帧速率时间混淆，不同的物理现象，如科幻流明的股票，例如，[面包车Roosmalen1999年不规则的老化; Piti'e等。2004年; Piti' E等人。2006年;德隆和Desolneux2010; Farbman和Lischinski2011; RE：2015年远景规划]。这些方法是互补的，以他们的工作。他们表现好这些任务，但要如灰度降解[Piti'e等特定领域的假设。 2004年]，全球影响[Farbman和Lischinski2011]，或称为退化模型[Piti'e等。 2006]限制其适用于其他的应用程序。另一方面，他们的方法着重于在其中由图像音响产生的时间伪像滤波器施加于每个帧独立地为不同的情景。输入视频本身的不稳定性不属于他们的工作。他们的做法是在精神梯度域图像处理技术，其目的是保持高频场景内容，同时允许对已成功为无缝合成[P'erez等其它目标的方法低频调节相似。2003]，HDR压缩[法塔勒等。 2002]等几个滤池，包括视频解块[Bhat等人2010。他们的工作不同于这些技术中，它着重于视频中的时间的一致性和是不可知的所施加的滤波器。



1. **恢复时间一致性**

他们的算法的输入是未经处理的视频{V0，V1...}，并且在同一视频{P0，P1，...} = {F（V 0），F（V 1）...}处理一帧一帧由一些图像滤波器F。该滤波器F已经出台，他们设法消除创建一个暂时稳定的视频{O0，O1...}时空的文物。在空间上，工件P中可以是全局或局部。例如，固有图像分解是德音响定义最多到全局乘法因子和算法通常设置该因子任意，导致随机偏移在每帧中。依赖于复杂的优化方案的算法很容易发生局部最小值，这使得它们过于敏感，初始条件，并且可以产生在相邻帧之间不连续的局部变化。此外，许多优化方案在空间上是转正，所以变化通常会影响整个对象或图像区域的一次 - 他们很少出现在几个像素的规模。在时间域中，这些工件的廓是任意的：它们可以变化缓慢，是随机在每帧，或在之间的任何地方。他们与这些特点考虑设计他们的算法。

一个简单的方法是通过光学溢流扭曲的第一个框架来重新创建每个后续帧执行完美的时间一致性。然而，这忽略了累积溢流网络的视场不可避免的不精确性，并最终造成较大的误差。此外，该方案没有考虑像闭塞和外观的变化，例如，由于光线变化的问题。换句话说，执行时间一致性可以来在现场动力学为代价。他们的解决方案平衡了这两个方面。

3.1 时间一致性和场景的动态联合优化

他们描述了他们的方法在因果设置：他们考虑第n帧（n>1），假设先前帧已被处理。这个处理帧一次一个，这使内存要求小，能够实现arbitrarily-长视频处理而不必借助复杂的内存管理方案。

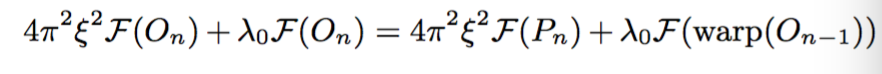
他们制定用一个简单的函数：../Library/Containers/com.tencent.qq/Data/Library/Application%20Support/QQ/Users/444871403/QQ/Temp.db/282C49FD-02EA-43F1-9E76-66907FC85F74.png，其中x表示在帧中的空间位置，和warp（）使用后向流动到平流输送前一帧向着所述当前帧。

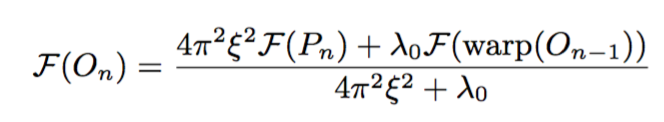
对于现场的动态来看，会有一个简单的选择：arg minOn ∥On −Pn ∥ dx然而，P在时间收到的影响不一致，所以这个会不利于他们的目标和视频传输不稳定。理想地，他们想与O只附加至P的，表示场景的一部分，并丢弃的那部分不一致的。他们从前辈那边获得灵感，通过

../Library/Containers/com.tencent.qq/Data/Library/Application%20Support/QQ/Users/444871403/QQ/Temp.db/32646ADF-6B02-4BED-80CE-6506D40E02F1.png直观地，这可以看作是一个数据连接上的场景的边缘，其中所述梯度的边缘的近似。

3.2 频域分析

他们的方法具有在不同空间中频率变化的影响，在视频中低频率被更多的限制为比高频时间上一致。能得出公式：

 进而推导出

，这是他们最基本的分析。首先，他们来看看的作用，对于大的值，即 他们有也就是说，他们总结先前帧这为他们前面所讨论的时间一致性，对于动态场景无效。对于小值，以致 可得到也就是说，他们复制由图像滤波处理的帧的频率成分不附加任何时间一致性。DC分量（）是通过不同的处理。如果， 他们会得到 0/0 对应于公知的不适定性泊松方程在没有边界条件的不确定性. 更有趣的是，如果 他们会得到 他们复制前一帧的DC分量。即使具有小时间加权，只要它是不为零，他们的方法消除了闪烁由于一个恒定的空间偏移。

接下来，他们为0的一般非零值和分析不稳定滤波后的帧的相，以对比德的影响稳定的前一帧其结果是，稳定化的帧在被占据着前一帧上的低频率，但高频率更接近于图像滤波器的输出光合的。这意味着该时间一致性更强烈地强制上低频。他们说明上的合成，回想一下，他们的目标是去除时间不一致性大多恒定或大规模同时保留由不年龄边缘捕获的场景结构这个属性。通过更强烈的低频执行时间一致性和维护高频率，他们的做法符合他们的目标。

**3. 实施**

warp()算子方程1a的选择是关键的，因为在输入视频V不精确对应可能导致闪烁，或在稳定的结果O.测试几个光流技术后渗色，他们发现，该方法在产生令人满意的结果的视频，尽管偶尔会引入轻微渗色。其主要缺点是计算成本，以1-2小时为100帧在1024 X 576分辨率。他们也考虑了几个近邻领域的技术。 PatchMatch提供一个互补的选项，产生令人满意的结果在许多例子，包括对视频哪个是具有挑战性的光流，并且在成本的一小部分：小于30秒100帧。然而，PatchMatch有时引入了轻微的闪烁当估计的对应字段在不连续的两个连续帧之间。在一般情况下，这两种方法都能够生产出高品质的结果，尽管该方法实现最稳定的输出取决于特定视频和应用：在快速运动的视频往往更好地与PatchMatch处理，而应用其中除去纹理线索，如深度预测和内在的分解与光流较好。

**4. 小结**

上文已经描述了他们的算法，用来稳定的施加逐步到视频图像过滤器的输出，依赖线性系统从而来解决标准最小二乘。通过分析在傅立叶域中的方案的属性，同时表明，虽然表明上简单，但是执行高频和低频之间的分化很复杂，从而来使视频变得稳定，从而不会降低其含量。他们的实验结果表明了，技术的性能比平滑的时间效果更好，同时能在个钟独立于它们的内部工作的应用程序中产生高品质的结果，来帮助视频处理工具处理该图像。

参考文献

[1]Blind Video Temporal Consistency Nicolas Bonneel,James Tompkin,Kalyan Sunkavalli et al.

[2]AUBRY， M.， PARIS， S.， HASINOFF， S.， KAUTZ， J.， AND DURAND， F. 2014. Fast local Laplacian filters: Theory and applications. *ACM Trans. on Graphics (SIGGRAPH)*.

AYDIN， T. O.， STEFANOSKI， N.， CROCI， S.， GROSS， M.， AND SMOLIC， A. 2014. Temporally coherent local tone mapping of hdr video. *ACM Trans. Graph. 33*， 6 (Nov.)， 196:1–196:13.

[3]BARNES， C.， SHECHTMAN， E.， FINKELSTEIN， A.， AND GOLD- MAN， D. B. 2009. PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. on Graphics (SIGGRAPH) 28*， 3.

[4]BELL， S.， BALA， K.， AND SNAVELY， N. 2014. Intrinsic images in the wild. *ACM Trans. on Graphics (SIGGRAPH) 33*， 4.

[5]BESSE， F.， ROTHER， C.， FITZGIBBON， A.， AND KAUTZ， J. 2012. PMBP: PatchMatch belief propagation for correspondence field estimation. In *BMVC - Best Industrial Impact Prize award*.

[6]BHAT， P.， CURLESS， B.， COHEN， M.， AND ZITNICK， C. L. 2008. Fourier analysis of the 2D screened Poisson equation for gradient domain problems. In *ECCV*， 114–128.

[7]BHAT， P.， ZITNICK， C. L.， COHEN， M.， AND CURLESS， B. 2010. GradientShop: A gradient-domain optimization framework for image and video filtering. *ACM Trans Graph (SIGGRAPH) 29*， 2.

[8]BONNEEL， N.， SUNKAVALLI， K.， PARIS， S.， AND PFISTER， H. 2013. Example-based video color grading. *ACM Trans. on Graphics (SIGGRAPH) 32*， 4.

BONNEEL， N.， SUNKAVALLI， K.， TOMPKIN， J.， SUN， D.， PARIS， S.， AND PFISTER， H. 2014. Interactive intrinsic video editing. *ACM Trans. on Graphics (SIGGRAPH Asia) 33*， 6.