 硕 士 研 究 生 读 书 报 告



题 目 盲视频的时间一致性

作者姓名 李 旦

作者学号 NB15221

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

Blind Video Temporal Consistency

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

Li Dan

Zhejiang University， P.R. China

2015

**摘要**

人们在拍摄事物的时候，为了使得拍摄的事物更加美观，通常会给自己拍摄的事物加上自己喜欢的滤镜。滤镜的出现使得该团队的照片与视频更加丰富多彩。而滤镜属于图像扩展的一部分。虽然扩展图像处理在对照片的处理时候是一项简单的技术，但是对拍摄视频来说却不是一项简单的技术。在拍摄视频时，增加图像扩展处理会独立处理视频的每一帧，在这种情况下，通常会导致视频空间和时间不一致。那么如何解决造成这种视频扩展图像处理的时空不一致的问题，已经成为一个新的研究方向。为了解决这个问题，各个团队都提出了自己独创的算法和处理方法，每个算法和处理方法都有各自的特点。本文论述了关于一个团队是如何对视频扩展图像处理的时空不一致做出相应的处理，这个方法的优缺点是什么，以及最终的效果如何。

**关键词：**视频扩展，时空不一致，处理算法，处理方法

**Abstract**

People at the time of shooting things，in order to make more beautiful things， we will give own shooting things with some filters. Filters the advent of our photos and videos more colorful. The filters are part of the image extension. Although image processing in the handling of the photo is a simple technology， for shooting videos， is not a simple technology. When shooting videos， increase image extension processing will handle each frame of video， in this case， usually leads to videos temporal inconsistency. So how to solve this video extended image processing of temporal problem， has become a new research direction. Every team in order to solve this problem， put forward their original algorithm and processing method， each algorithm has its own characteristics and processing method. This paper discusses about how the method for video image processing expanding temporal inconsistency makes the corresponding processing， what is the advantages and disadvantages of these methods， and how the effect of treatment.

**keywords:**Videos Extension， temporal， algorithm， process method

**1 简介**

先进的处理，有效的重构过滤器，创造性、高级的编辑和分析是对当前静态图片。另一方面，视频却没有相同各式各样的工具。对视频的处理只有一种简单的对用静态图片处理方式用来对视频的每一帧做独立处理。这可能在一些简单的情况下像高通和低通滤波有效，但在许多其他更复杂的情况下，这种方式将会产生难看的结果——闪烁。产生这样的原因是多方面的，例如:一个复杂的优化技术分落在不同取决于帧的最小值的间隔中，或者一个过滤器可能依靠统计学，像平均颜色一样，整个视频序列却不稳定。

解决这个问题的一个方式是显式地占时间一致性。沿着这条线开发了几个视频处理算法，如颜色分级 ，动态范围压缩 ，本质分解和色调稳定。这些技术虽然有效，但这都只能运用于特定的任务，不能延伸解决其他问题。Paris 和 Lang et al提出更通用的方法来扩展静态图像方法用于视频。然而，假设一个特定的过滤器配方，这却限制了他们的应用程序。Lang et al等人求助于低通过滤器。可以见得，这种方法降低了闪烁但却无法去除。

本文介绍的团队的目标是更一般的方式通过扩展图像的方式运用在视频上，并提出一种是不知内部过滤器设置的算法，例如：他们将图像过滤器作为黑盒，用输入帧生成处理帧。在这个意义上，他们的方法是接近忽视被应用于图像过滤器的时间一致性。有两个要求：1、原视频可用光学流或特性是可恢复的，这样，它可以用作时间一致性。2、过滤器不产生新的不相关的输入内容——例如，绘画滤器顺序生成画笔描边的纹理和绘画技术，合成新的内容就属于超出范围。也就是说，该团队的方法涵盖了各种各样的过滤器，如去雾自动摄影提高，颜色分级和内在的分解。他们制定梯度域算法，提出一个能量函数，相当于一个空间筛选与时间约束的泊松方程，可以有效地解决时空不一致。他们描述他们的方法的属性根据频率的内容。他们提供实验证明，该算法可以处理各种各样的影响，而且它可以直接扩展最先进的视频图像过滤器。

他们的贡献

•由于视频帧图像滤波器的应用消除闪烁技术。

•一个梯度域形成的时间一致性问题无关应用图像过滤器。

•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**

逐帧应用的时候一些图像过滤器像低通滤波过滤器等图像过滤器及其边缘保护共同产生暂时稳定的结果，然而，许多其他图像过滤器无法应用到视频上，而且需要适配。

一个流行的解决方案是专注于一个给定的滤波器为一个特定的应用程序。例如：传送一个视频到另一个视频通过的暂时的颜色过滤器。让视频基调映射通过将HDR内容分解为基础层和细节层暂时一致。通过自身的分解生成稳定的视频。所有这些方法只适合他们自身的需求，但是因为这方法是他们依靠于他们详细的目标应用程序而强制时间一致性，并不适用于其他应用。相比之下，该团队提出一个方法，适用于大量的图像过滤器。

更多普遍的用来处理全类过滤器的方法已经被提出。通过扩展高斯内核延伸到时间域，并使用这个结果适应双边滤波和均值漂移聚类等应用视频。通过将时域平滑的概念扩展到利用光流和基于运动估计技术和彩色化的重新优化。对于其他不优化能源的技术，他们求助于时间的低通滤波。该团队将看到在结果部分时间高频闪烁平滑降低，但低频依然不稳定。与这两种方法不同的是，该团队的方法不需要图像过滤器有一个特殊形式也不用适配它的形式。相反，该团队运行没有修改的过滤器作为一个黑盒并创建一组过滤框架，然后处理他们的清除时间工件。

并发工作，现在技术处理未知的图像过滤稳定视频帧，可以表达为非线性曲线应用于视频帧原始的区域。相比之下，该团队的技术并不局限于一个特定的配方和能够处理应用内在图像或深度。

并行技术已经发展到可以移除不一致的输入视频，而不仅仅是一个视频处理过滤器。从捕获时间等问题机内自动白平衡到暂时混淆不匹配的相机/照明/投影仪帧速率，到不规则的老化的电影物理现象。这些方法对该团队的工作是互补的。他们执行这些任务，假设特定的领域灰度退化等或已知的退化模型限制其他应用程序的适用性。另一方面，该团队的方法是在一个不同的场景，在该场景中，时间构件都是由一个图像过滤器生成应用于每一独立帧。忽视那些本身不稳定的输入视频。

该团队的方法是相似的梯度域图像处理技术，旨在保护高频场景内容，同时允许低频向另一个目标无缝合成的方法调整，已成功HDR压缩和其他几个过滤器包括视频解封。该团队的技术与这些技术的不同之处在于，它专注于视频时序一致性和不用知道具体的应用过滤器。

**3 重构时空一致性**

该团队的算法输入一些的视频{V0，V1……} 和相同的视频{P0，P1 ...} = {f(V0)，f(V1)...} 通过一些图像过滤器逐帧处理 f。过滤器 f 引入了时间工件以至于该团队试图移除创建暂时稳定的视频{O0， O1 . . . }.

空间上，在P内的工件即是全局也可以是局部的。举个例子：空间内部的图像全部分解然后定义为全局乘法银子和算法常常设置这个因素，然后随机偏移到每一帧。依靠先进的优化方案的算法很容易引起局部极小，这使得它们对初始条件过于敏感，可以生成相邻帧之间的不连续的局部变化。此外，许多优化方案在空间上是发生变化，所以变化通常会立刻影响他们，很少发生在几个像素的规模整个对象或图像区域。在时间域中，这些工件的轮廓是任意的：它们可以缓慢的变化，随机在每帧，或在之间的任何地方。该团队基于这些特点设计他们的算法。

一种简单的方法是通过光流的第一帧来重新创建每个后续帧执行完美时间一致性。然而，这忽略积累流域不可避免的会有不精确性，并且最终将导致较大的误差。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牺牲动态场景为代价。该团队的解决方案平衡了这两个方面。

**3.1联合优化的时间一致性和动态场景**

该团队描述了该团队在因果设置的方法：该团队考虑第n帧（n>1），假设先前帧已被处理。这帧处理一次，这使要求内存需求小并且能够实现任意长的视频处理而不必借助复杂的内存管理方案

该团队制定一个简单的最小二乘时间一致性的目标： 其中x表示在帧中的空间位置，并且向着所述的当前帧使用回流动到平流输送前一帧。

从动态的场景来看，人们选择将是一个简单的数据连接词 . 然而，P隐式地从时间上不一致受到影响，所以这个词会违背该团队的目标和转换成不稳定输出视频O.理想上来说，该团队只想O附加至P，表示场景的一部分，并丢弃不一致的那部分。要做到这一点，该团队从Elder发现场景是其边缘代表中汲取灵感。该团队也受泊松图像编辑的启发，它通过复制其梯度再现图像区域的外观。而不是要求是相似的像素值。因此，该团队获得了他们的梯度是相似，该团队得出最小值：。直观的说，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。

该团队结合两个术语调制对时间一致性项的重量影响，w(x)，其输入视频在V后面测量一致性。该团队设置的第一帧作为边界条件和计算O为最小的最小二乘能量：

 (1a)

 (1b)

 (1c)

加权函数w(x) 是关键，该团队的方法，因为它可松弛时间一致性要求当输入视频V本身并不一致或warp()不准确，这个时候该团队检测warp前一帧无法很好的解析当前帧（等式1b）中。换句话说，该团队只强加时间一致性，他们使用的第一帧作为对应于常见的情况，其中用户将图像调整滤波器f以实现第一帧上所期望的结果，并希望传播的边界条件（1c）的结果以相同的质量传输到视频的其余部分。这将是直截了当地调整该团队的方案是使用任何帧为基准，从中处理视频向前和向后的时间。该团队的配方通过参数化，它表明该团队有信心弯曲控制正规化强度。

为了使得方程式1a最小化，该团队使用了欧拉 - 拉格朗日公式推导得出的一个必须满足在最小的差分属性：

 (2)

其中在右侧的，和的所有数量是已知的。这个方程被称为筛选泊松方程;它是标准泊松方程与零阶项的一个变形。存在有效的方案来解决它。该团队的方法适用于大量的应用和不需要应用于一个特制的过滤器的每个帧。

**3.2频域分析**

该团队的方法具有在不同空间频率的变化的影响，在视频，与低频率被更多的限制为比高频时间上是一致的。该团队分析方程2在用，用于傅里叶变换和ξ为空间频率的频域。为简单起见，并在本节只，该团队假设加权函数w对于所有的x是常数并等于。即 应用这个恒等式  与等式2可以得到

 (2)

转换成  (3)

这是该团队的分析的基础。

首先，该团队来看看的作用，对于大的值，即 该团队有也就是说，该团队总结先前帧这为该团队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，该团队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， 该团队会得到 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 该团队会得到 该团队复制前一帧的DC分量。即使具有小时间加权，只要它是不为零，该团队的方法消除了闪烁由于一个恒定的空间偏移。

接下来，该团队为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。该团队说明上的合成，回想一下，该团队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，我们的做法符合该团队的目标。

相比较而言，时间平滑可以被建模为平均值老化的当前帧与前一帧，例如 也就是其中控制平滑强度。这将直接转化为。所有频率由前一帧同样受到影响，也就是说，时间一致性是执行均匀的频谱。

该团队分析该团队的时间一致性算法的时间噪声与合成实验中不同尺度的行为。该团队使用来自Sintel序列的46帧片段 —— 一个综合渲染动画与地面真实反射和阴影。该团队产生随机噪声，把它添加到一个随机每帧偏移和空间滤波，它们与高斯滤波器具有不同的标准偏差，以产生具有不同特征尺度噪声。

该团队加入该噪声层到地面实况反射率序列模拟由一帧一帧的处理，且在时间上的文物。该团队扩展噪声层，以保持恶化的反射率固定为25分贝所有噪声尺度的PSNR。该团队模拟“根据需要”的第一帧上通过加入无噪声到其设置过滤器的普通用户的做法。该团队计算上使用的技术中的原始序列的光流。该团队的时间一致性技术的性能作为添加的噪声的比例被改变。

在最小刻度，该团队的技术通过提高反射率的重建大约4.5分贝 - 这在很大程度上是校正随机每帧偏移的结果。作为噪声增大规模，该团队的技术确实更好，至多约33分贝一个PSNR。该团队也比较了郎咸平等人的时间平滑技术。 该团队还增加了另一个无噪声帧(30帧)之后，以模拟用户定期检查的输出的质量。这对该团队的做法影响不大，但有助于郎等人的技术，通过提供它的参考点向后传播。

高频噪声（< 4个象素），时间平滑能执行更好，但，随着噪声增加的比例（ > 4像素），该团队的方法变得更好。这表明，时间平滑更适合用于过滤高频扰动，例如传感器噪声，但是，该团队的方法是，在处理中，低频率的误差像不稳定不年龄滤波器引入的那些更好。该团队显示在验证部分，该团队获得现实世界的视频相似的结果。

**4结果**

本节提供了实施细则，与国家的最先进的技术比较，该团队的技术应用到多个视频处理应用。该团队的时间一致性，可以在所附的补充视频，其中还包括附加的结果更好地观察。

**4.1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，该团队发现，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。该团队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。

**4.2比较**

该团队比较该团队的方法来不变的滤波算法对典型的视频处理任务自动提高使用Adobe Photoshop的“自动颜色”，“自动对比度”的组合的视频帧的颜色和色调，而“自动色调”的工具。在每帧的基础上产生强高频闪烁和颜色缓慢漂移应用这些工具。如果无法访问原来的增强算法，郎等人的做法相当于时间平滑。如图4，采用平滑在他们的算法不能完全除去的时间抖动，这符合该团队的分析合成数据的值低。另一方面，使用更强的空间平滑导致不良的空间模糊，这是该团队推测是由于光流的错误。平心而论，郎等人。专注于重新制定的能源的过滤器，并建议该团队用在这里只是作为后备的时间平滑。所有序列的比较可以在补充材料中找到。该团队也比较适用现有的视频增强工具的Adobe Premiere的“自动颜色’增强了同样的场景。该团队启用了“时间平滑”选项，并获得了不稳定的结果，即使有大的平滑窗口。相比之下，该团队的方法不产生空间细节的损失在时间上一致的输出。

**4.3应用程序**

该团队的技术的优势在于，它使有关应用于视频帧的底层图像处理很少的假设事实，该团队已经探索将它应用到广泛的应用。特别是，该团队确定了在这些应用中颞不一致几个来源。高时间频率闪烁往往被敏感的参数，问题不适定性，或快速场景内容的变化引起的。低时间频率的文物通常是由在场景内容平滑变化引起的。该团队的算法妥善处理好这两种器物。

**4.4讨论**

该团队的方法解决了应用不稳定图像过滤器将视频引入时间不稳定的问题。它的目的是利用一对未处理的处理的视频和不意味着处理单个序列的场景。例如，它不能因在捕捉时的问题，例如传感器噪声或缩时的视频时间混淆消除抖动，因为，在这些情况下，不存在时间上一致的输入视频被用作参考。

该团队发现，该团队的方法不上铺垫工作得很好，因为出现模糊对象边界的不稳定性，这也是在那里光流技术往往失败。此外，如前面所讨论的，用于创建内容不相关的输入的艺术滤镜也存在问题，如绘画的风格化。

**5 结论**

我们已经描述了该团队的盲算法，以稳定的施加逐帧到视频图像处理滤波器的输出。该团队的方法依赖于能够用线性系统解决的标准最小二乘。该团队已经分析了在傅立叶域中方案的属性，并表明，尽管它的表面上的简单，它执行高频和低频之间的复杂的分化，使视频的稳定，而不会降低其含量。该团队的实验结果表明，该团队的技术的性能比时间平滑显著更好，并能生产出高品质的结果在各种各样的独立于它们的内部工作的应用程序，从而有助于使视频处理工具箱接近奇偶与该图像。

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.