

指导教师： 杨涛 提交时间： 2016/3/13

CVPR2015 Paper Translation

No: 01

姓名： 王博伦

学号： 2013302603

班号： 10011306

较大动作幅度下的影像放大

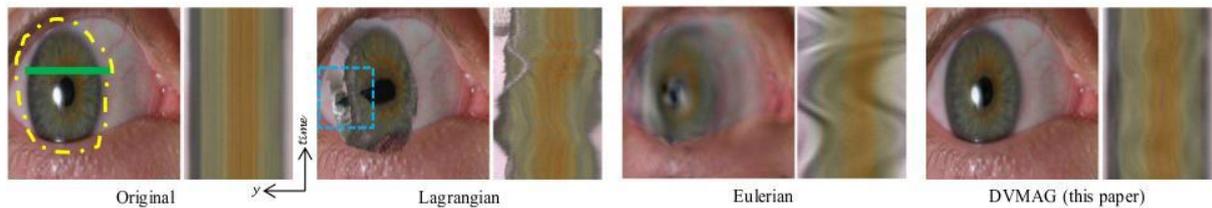


Figure 1:

摘要

影像放大技术展示了肉眼不可见的微小变化。为达到这个目的，当前的技术水平要求影像中的动作幅度非常小，但是不幸的是现实中这个条件通常并不能得到满足。微小但是有价值的变化经常会和较大的动作幅度结合起来，比如说当一扇门在转动的时候的微小变化，或者是运动中的眼球扫视到的视觉范围。在本文中，我们提出了一种基于层次模型的影像放大的方法，这种方法能够在巨大的形变之中方法微小的变化。经过测试的区域将会被临时对齐，借此微小的变化就可以被放大。我们用消光技术来放大那些我们感兴趣的区域，同时维持其余区域的特性不变。实验结果表明，当处理幅度较大的动作时，我们可以得到更高的放大率和更低的失真率。

1.简介

客观世界充满了各种各样的人眼不可识别的微小变化。比如，血液循环的时候皮肤颜色会发生变化，有风的时候建筑物会有轻微的摇摆，再比如，每一次心跳都会造成头的摇晃。虽然这些变化太小我们不能直接去观察，但是我们可以通过计算手段对其进行放大，这样就可以实现一个微小形变组成的世界，这是很令人兴奋的。当前的影像放大方法假设我们感兴趣的区域形变量非常小。然

而，许多我们感兴趣的形变都是比较显著的。当我们的身体在做较大的动作时，我们的皮肤事实上发生了微妙的形变。关闭的收费杆形变事实上很小，尽管它旋转的幅度很大。扫视面积通常是很大范围的眼睛移动相关的（见 Fig 1）。另外，影像或者物理也许是被手持相机捕捉的，又或者并没有一直静止不动，这样的话通常的影像放大方法不仅会放大我们感兴趣的区域，同时也会放大抖动的程度。如果把现在的影像放大技术应用到那些包含较大幅度运动的影像中，就会造成比较大的误差，比如光晕和影像错位，即使这样，那些小幅度运动影像还是很难去观测，因为那些微小的形变被大幅度的误差所掩盖了（见 Fig 1）。

在特殊情况下，也许有可能通过在放大步骤开始之前通过影像的稳定化预处理来去除不需要的失真，但是这种方法不适用于通常情况下的物体运动，就算是相机的抖动也不行。人们必须小心处理，因为任何微小的稳定化误差都会在放大步骤中被放大。这个问题在运动物体的边缘部分表现的尤为突出，比如一支胳膊和它之后多动作融合的背景，背景中可能会包含将要巨大的形变，需要被放大的微小形变，以及背景动作。现有的方法，包括线性 Eulerian 方法，基于相位的方法以及 Riesz 算法都假设存在一个单一的动作。上面这几种方法都确定了物体边缘的背景延迟效应，另外 Lagrangian 方法能够检测动作错误，因此可以用来确认放大的噪声。

本文提出了一种能够在较大形变存在的基础上进行微小形变影像放大的方法。

我们称之为 DVMAG,或者叫动态影像形变放大技术。用户选择一个他们感兴趣的想要放大的区域。我们就可以通过正规化参数模型来压缩该区域的显著形变。我们使用层次分解和叠加的方法来处理运动部分和背景之间的边缘部分。我们发现,同时将 Eulerian 影像放大方法应用到前景和它的叠加层之后,我们可以动态地降低失真率,在此之后,我们利用组织合成方法处理了放大后的形变,并且处理了潜在的闭塞问题。

本文的重点部分在于: 1)通过稳定化艺术在放大的过程中处理掉了大形变的影响, 2)使用放大区域的消光技术来保持正规化。3)在 Eulerian 放大方法中加入组织合成方法来填充显著的空洞部分。

2. 艺术描述

影像放大技术要求在图像序列中放大并且虚拟化微小的形变。现在的方法主要被分为两大类: Lagrangian 方法和 Eulerian 方法。在 Lagrangian 方法中,需要精确地预测动作,这里的动作所指的就是那些需要被放大的形变。相反的,在 Eulerian 方法中并不需要精确地预测动作,而是通过计算非动作部分的补偿帧之间的差异来预测形变。Lagrangian 方法只能放大动作的变化,而 Eulerian 方法除此之外,还可以放大颜色的变化。

Liu et al.较早地提出了影像放大技术。特征点的轨迹被分解,并且被分割成静态和动态两个集合。仿射行为模型能够适应静态模型,这种模型在参考帧上记录那些已经被测试过的帧序列。行为将会被多次预测,放大并且会被加到已记录的序列中。这种方法产生放大后的输出。

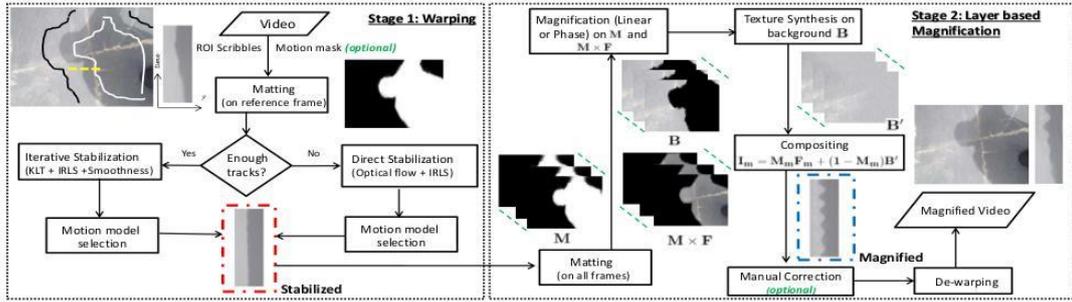
Eulerian 影像放大技术提出了一系列技术,在文献[17]中,一个输入序列首先会被分解成为一个多尺度的栈(Laplacian 或 Gaussian)。微小的变化将会被暂时过滤,经过尺度变换并且加回到输入序列,就会产生放大后的序列。经过这个步骤,

产生了一些有意义的结果,但是我们只能处理那些小的形变和被放大的元素。更大的形变将会产生裁剪失真,这些失真的结果将会掩盖放大的影像中有用的信息(见 Fig 6)。第二种 Eulerian 影像放大技术基于对相位的观测,这种方法会测量那些与动作相关的形变。输入影像被分解为多尺度,多方位的栈。每一组帧的幅值和相位都是分离的,在位置,方位和尺度上,帧的相位将会被临时过滤。这种方法可以预测图片序列中微小的临时变化。与之相对应的相角变化将会被放大并且重新加回到输入影像中。再重构时间域和空间域,就会得到输出的放大之后的影像。基于相位的技术在控制噪声方面比线性技术拥有更好的性能。然而基于相位的技术的最大缺点在于其依然不能处理大幅度的动作。如果强行用于处理大幅度动作,将会产生显著的模糊的失真(见 Fig 1)。

我们所做的工作和 Bai et al 的工作是相关的,也是一种选择策略,由用户选择自己想要进行稳定化的区域。然而这种方法会在之前那些动作变化比较少的区域带来不自然的变化。一种基于图像分割的优化方案就被用来将固定的区域与静止的帧进行融合。他们的稳定化阶段和我们的有关联,但是我们提出了一个完全不同的问题 - 影像放大。以及需要不通处理的不同形式的失真度。

3. DVMAG: 动态影像放大

我们提出了一种在较大动作幅度影响下的放大微小动作的影像放大技术。我们的技术(见 Fig 2)有两个主要部分: 1.对影像进行弯折来减小较大幅度动作的影响。2.基于层次模型的放大。在弯折阶段我们在搜索去除较大幅度动作的同时保留较小幅度的动作,这样做我们就不会引入被放大的失真。在这个阶段中,我们使用 KLT 跟踪算法,可见流或者是正规化的降序参数模型来处理较大幅度动作,基于层次模



型的放大方法会根据首次重叠将一副图像分割成为前景和背景。之后，我们放大每一层，就会倒序输出一个放大后的图像序列。我们使用组织合成的方法来填充结果图像中的空洞。最后，我们将放大后的序列结算成原始的坐标图。用户通过标记在参考帧上指定自己想要放大的区域（见 Fig 2 的左上）。

3.1 弯折

给定一个输入序列 I ，我们想要预测稳定化的序列 $I(s)$ ，为此，我们可以在一个参考帧里临时记录该序列。我们可以用降序的模型(fai)对 ROI 里的大幅度动作进行建模来保留那些需要放大的小幅度动作。

$$I^S(x, t) = I(\Phi_{r,t}(x), t),$$

公式中的 x 表示二维点坐标。给定一个第 t 帧的点集 X_t 和与之相关的 X_r ，通过求 $\|\Phi_{r,t}(X_t) - X_r\|_2$ 的最小值我们得到了 $\Phi_{r,t}$ ，我们使用两种方法来得到 X_t 和与之相关的 X_r 。一种是 KLT 跟踪法，一种是可见流方法。

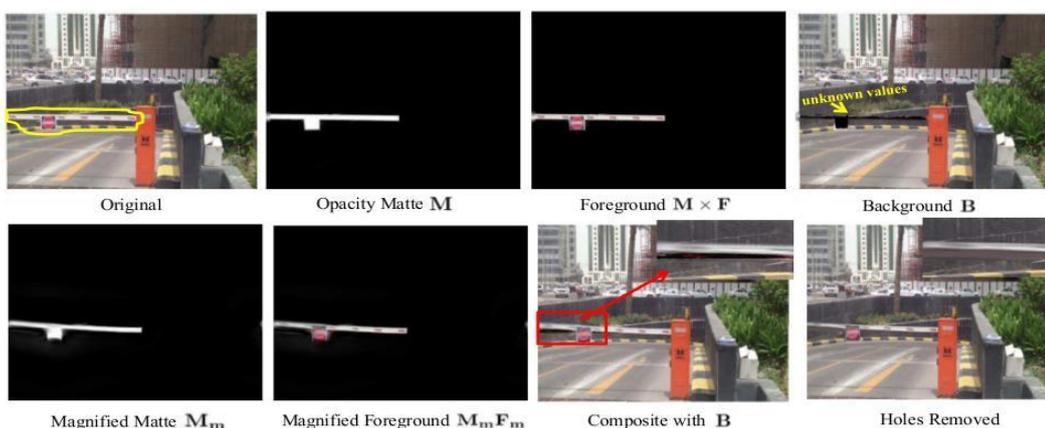
KLT 跟踪中的预测 $\Phi_{r,t}$ 流程：对于第 t 帧中的每一个点集以及与之相关的下一帧的点集。通过仿射或者变换模型来匹配两者。利用移动均值滤波器向 Φ 添加普通的平滑。另外，我们使用以测试帧 t 为中心的五帧窗和有权窗来降低匹配错误率，这些错误的产生通常是由于我们在求

Φ 的时候使用了 IRLS 方法。这里所说的权值反比于 $\|\Phi_{r,t}(x) - x\|_2$ 。我们在这里使用一个临时的迭代方案来预测待测试帧 t 和参考帧之间的 Φ 。首先，对于每一对连续帧我们得到一个 Φ 的预测值。比如说，对于 $r > t$ ，我们就可以依次得到 $\Phi_{t+1,t}$ ， $\Phi_{t+2,t+1}$ ， $\Phi_{t+3,t+2}$ ，..... $\Phi_{r,r-1}$ 。因此从 t 到 r 的直接变换就变成了 $\Phi_{r,t} = u_{r-1}$

$\Phi_{u+1,u}$ 。给定一个 $\Phi_{r,t}$ ，就可以利用等式 1 来稳定化第 t 帧。对于 $r < t$ ，我们进行同样的操作，但是时间域上是相反的。这样就可以在参考帧 r 上得到整个帧序列 I 的稳定化结果。

利用可见流方法进行 $\Phi_{r,t}$ 的预测：相比于特征点的轨迹，可见流方法对于动作误差更为敏感，因此就很有必要进行调整：我们不使用临时的迭代方案来预测 $\Phi_{r,t}$ ，因为误差是可以积累的。因此，我们应该在待测试帧 t 和参考帧 r 之间直接预测 Φ 。在这里，我们在参考帧 r 和其他所有的待测试帧序列之间预测可见流，方法在文献[6]中有所陈述。考虑完这个因素，我们就可以为特征和点的轨迹继续预测模型的参数。

可见流与特征点轨迹的对比：在弯折阶段，很重要的一点是在行为模型建立的时候选择良好的候选模型。这样就可以降低后续放大过程中的风险。因此，我们在 KLT 跟踪和可见流这两个方法中选择一种。我们只考虑那些隐藏行为的候选模型。首先计算 KLT 跟踪，如果追踪结果像素



点的百分比大于 KLT 跟踪法的阈值，否则的话，我们设定阈值为 5%，利用可见流的方法进行预测。

仿射模型与变换模型的对比：我们对 Φ 预测仿射和变换模型，并且找出平稳化误差最小的那个模型，

$$\sum_{t=1}^T \sum_{x=1}^P |I^S(x, t) - I(x, r)|.$$

$I(x, r)$ 就是参考帧， T 是平稳化序列 I^S 中的帧数， P 是一帧中的像素数。只有在用 ROI 方法进行计算的时候我们才会用这个公式。

3.2 基于层次模型的

Eulerian 放大方法

我们可以提出一个基于层次模型的影像放大方法。给定一个待放大区域，之后，将一副图像分割为三个层次：1.不透明区域。2.前景区域。3.背景区域。我们是用 Levin et al.方法堆图像进行去透明化处理，另外，我们用 Eulerian 方法对不透明区域和前景区域进行放大。如果我们想要放大图像中颜色的变化幅度，我们就可以使用 Linear 技术。否则就使用基于相位的技术。我们通过将放大后的前景图像置于原始的背景图像之上的方法来重构剩余的未放大

区域。另外，我们使用组织合成的方法来填充那些放大的前景图像带来的图像空洞。另外，如果需要的话，我们可以手动去除掉那些人为造成的误差。最后，我们将那些放大后的图像序列进行反弯折处理，并将其加入到原始的时空坐标中。对影像的每一帧都进行相同的处理。

Fig 3 更加详细地展示了一种用于放大停车场大门的方法(见黄色线条)。另外，去背景化技术会产生一些不符合 ROI 的误差。如果直接进行放大的话，就会产生巨大的误差(见 Fig 4 的左半部分)。因此，我们选择放大一个截取的前景图($M \times F$)而不是整幅图像(F)(见 Fig 4 的右半部分)。为了简单起见，我们将待放大图像表示为 $M_m \times F_m$ ，其中 M_m 是被放大的不透明图像。已知 M_m 和 $M_m \times F_m$ ，并且将被放大的前景图像放置到原始的背景图上，我们就可以得到一个综合的图像序列。然而，直接使用 B 将会在被放大的区域产生图像空洞(见 Fig 3 的下半部的红色部分)。因此，在进行图像的综合之前，我们预先通过组织合成的方法对未知的背景值进行填充(见 Fig 3 顶部的黄色文本)。给定黑色背景 B' ，我们通过公式 $I_m(x) = M_m(x)F_m(x) + (1 - M_m(x))B'(x)$ 来计算放大后的图像序列，公式中的 x 表示所有的图像像素点。我们对待测试的图像序列的每一帧都进行同样的操作，这样就能产生一个稳定的放大图像序列(见 Fig3 的最后一行的右边部分)。

Video	α	ω_l (Hz)	ω_h (Hz)	f_s (fps)
Eye	30	20	70	1000
Bulb	15	20	70	1000
Gun	20	8	33	480
Water	10	5	6	29
Shadow	50	2.8	3.2	29
Parking	40	4	5	29
Leaves	20	0.5	2	29
Sim1	0-120	4.9	5.1	24
Sim2	50	72	92	600

在那些前景和背景的像素值比较相似的区域会产生一些额外的误差(见 Fig 5 的 b)。为了彻底消除这一类的误差，我们允许用户选择性地抹掉一些特定的区域。用户只选择那些参考帧中存在误差的部分(见 Fig 5 的蓝色部分),之后整个影像部分中存在误差的部分就会被用原始图像序列的值替换掉(见 Fig 5 的 c)。通过实验我们发现，在我们实验的 9 个图像序列中只有一个图像序列需要进行这种手动的改正。

算法的最后一步将放大的图像序列 I_m 解码到原始影像的时空坐标中。在我们的实现方式中，我们是将之前保存的动作参数插入到被解码的图像序列中。

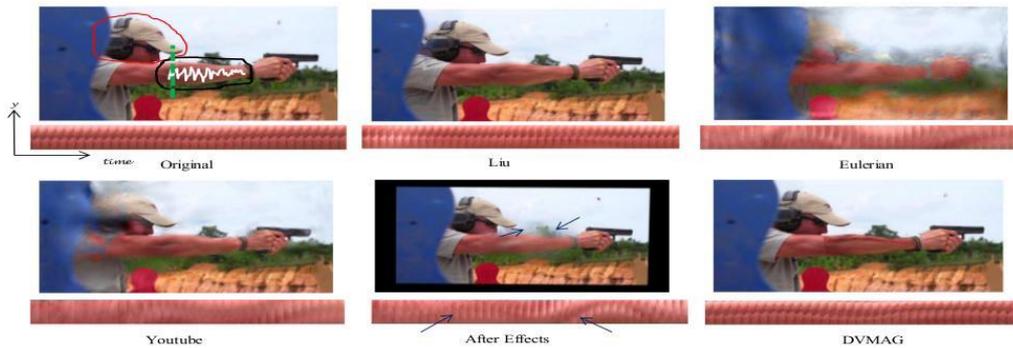
4.结果

我们用参考帧中的真实帧在原始图像序列和合成的图像序列上都做了相应的实验，最终我们像 Youtube 和 Adobe 那样，做了 Eulerian 放大操作。对于真实的影像序列，我们需要进行性能评估。对于可控的实验，我们需要针对 ground-truth 进行性能评估。结果显示针对小动作的艺术状态方法在处理大动作影像放大的时候也实现了优化。我们的 DVMAG 技术显著减少了误差，并且增强了域的可用性。Table1 展示了使用相应参数进行测试的序列信息。

4.1 艺术状态技术方法

我们将我们的 DVMAG 技术与其他两种影像放大方法，Lagrangian 方法和 Eulerian 方法进行比较，对于 Eulerian 方法，我们对 Bulb 序列使用线性技术，对其余部分使用相位技术。我们对线性和相位技术使用“作者”的实现方法，对于 Lagrangian 方法，对于原始影像数据来说，这个方法是不可实现的。对此，我们最开始使用我们自己的方法对影像进行稳定化处理，之后我们再使用 Liu et al.的可视流技术实现一个动作密集区域。为了评估我们的稳定化效果，我们将 DVMAG 方法

另外两种更深层次的放大方法进行比较，这里我们使用了两种离线的稳定化方法来对待测试序列进行稳定化处理。之后我们再使用 Eulerian 方法堆稳定化处理之后的序列进行放大处理。我们测试了 Youtube 和 Adobe 的稳定化方法。前者我们使用 Youtube 的视频管理器，对于后者，我们使用的是 After Effects 弯折稳定化工具 VFX。AE 允许用户定义一个动作掩码，这样的话对于平等的比较额而言，AE 的结果和 DVMAG 的结果使用的是同样的动作掩码。我们也对动作补偿技术进行了比较。我们使用可视流技术在每一帧及其参考帧之间产生动作预测。对于每一帧，我们使用密集的可视流来转移其像素。这样就可以产生待检测帧的一个预测。



对于 Lagrangian 方法来说，After Effects and Liu，我们只放大那些我们感兴趣的区域，对于剩余的部分，我们将整个影像帧进行放大，因为在许多情况下 ROI 是移动的，即使进行了稳定化操作之后也是如此。为了能够更加容易地进行评估，我们在 DVMAG 的参考帧上将所有的结果进行了临时稳定化处理，使用我们的技术，我们可以将 Eulerian 放大方法暂时进行放大。(见 3.1). 其余定义的技术应该先于放大操作。

4.2 真实的图像序列

Fig 6-8 展示了对于给定的图像序列，DVMAG 放大方法和其他不同的放大方法的比较。对于每一种技术，我们在放大后的帧上展示一个帧和一个时空片。在 Fig 6 中我们测试了 Bulb 图像序列。在这一序列中，一个人拿着一个灯泡并且将其在垂直方向上向上移动。用 DVMAG 方法进行处理的时候会出现临时的亮度变化。这种变化是由于电流的变化而产生的，并且除非将其放大，否则这种变化是很难被发现的。使用 Liu 的方法处理 Bulb 不会产生任何临时的变化。这是由于可视流的实现就是为了最小化变化。其余的技术也不会产生任何的临时变化。Eulerian 方法会产生一些颜色裁剪的误差(见标红的部分).这种误差的产生是由于为了过滤掉输入帧的错位。Youtube 的方法也会产生相

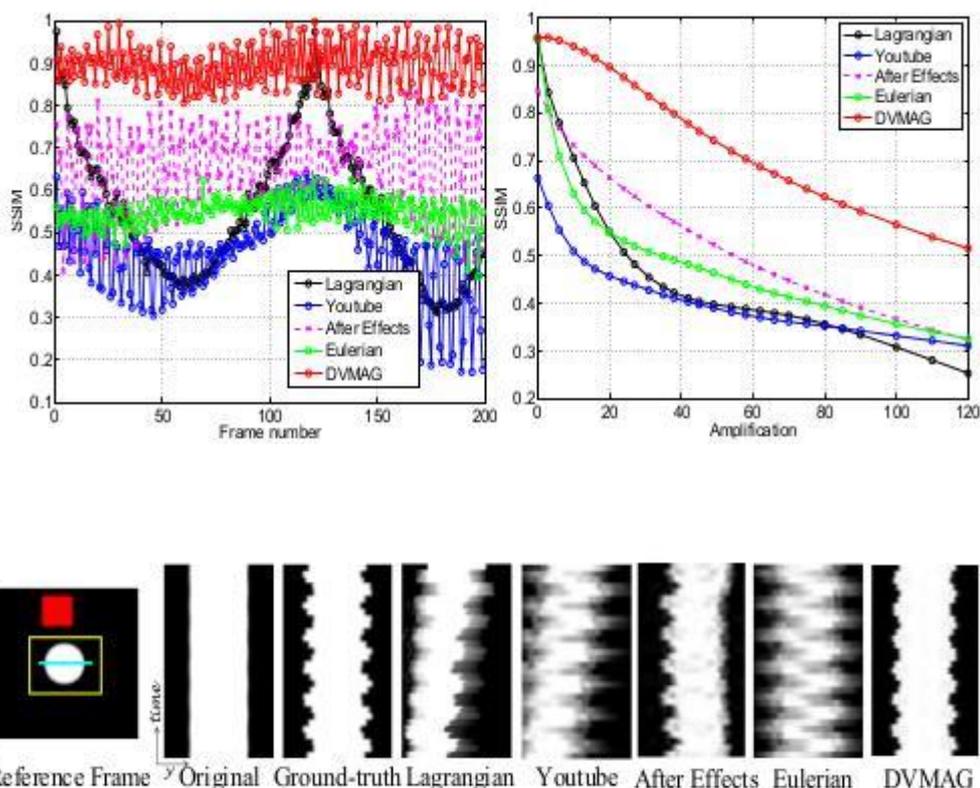
应的裁剪误差，最终 Lagrangian 方法就会产生噪声。

Fig 7 展示了对于停车场视频产生的结果。其中展示了地下停车场的入口。打开或者关闭停车杆都会造成大门的形变。在原始图像序列 DVMAG 方法对该序列进行放大，我们可以发现，人的胳膊在设计过程中实际上是在动的。Eulerian 方法，Youtube 方法和 After-Effects 方法都会产生模糊的误差，而 Liu 的方法根本就没有检测到胳膊的变化。

总之，我们的方法放大了我们感兴趣的区域，同时保持了临近区域的元素，并且比其他方法的效果更好。Eulerian 方法会产生模糊的结果，Youtube 方法的稳定化处理不能移除那些较大的动作因此也会产生模糊的结果。在 AE 的方法中，多个移动的物体会产生稳定化的误差，这些误差也会被放大。另外，AE 的方法通常都会对待测试物体的边缘区域产生一定的影像。Lagrangian 方法对动作的误差很敏感因此会产生噪声。直接的动作补偿方法(比如 Liu 的方法)几乎不能对微小的变化进行放大。这是因为可视流方法本来就需要将临时的变化降到最小。

4.3 可控实验

Sim1: 我们创建了一个参考帧，这个参考帧包括一个白色的圆还有一个红色的矩形(见 Fig 9 的左半部分).白色区域就



是待放大的 ROI，而红色的矩形被用来产生候选动作。我们定义一个局部动作 $d_j = A \sin(2\pi f f s j)$ ，在这个公式中， $A = 0.25$ 个像素点， $f = 5$ 周期/帧， $f s = 24$ 帧/S，我们通过水平方向上用 d_j 替换白色的圆来产生 Sim1 的帧 j 。对 200 帧图像进行同样的处理就会产生一个带有白色震动圆圈的序列。这里的 (A, f) 指的是形变量的幅值和相位。在此之后，我们使用 $\Delta j = A \sin(2\pi f f s j)$ 来对每一帧进行变换，这样就能将大的全局动作加到形变序列中。 Δj 是第 j 帧的全局动作， $A = 40$ ， $f = 0.1$ 。全局动作只发生在水平方向上，最终的结果序列就是 Sim1。

我们使用不同的影像放大方法堆 Sim1 进行处理。目标在于评估该方法放大白色圆圈的能力。我们测试了不同的放大系数 α 并且同 ground-truth 方法进行比较。在相同的方法处理 Sim1 的时候产生了 GT。对于放大系数 α 来说，通过使用公式 $d_j = A(\alpha + 1) \sin(2\pi f f s j)$ 对参考帧的白色圆圈进行相应的变换，我们就可以计算出相应的 GT。因为在 Sim1 中， j 代

表帧的索引， $A = 0.25$ 个像素点， $f = 5$ 周期/帧， $f s = 24$ 帧/S。对于 GT 来说，因为所有的比较都是基于临时对齐的放大方法，所以我们不需要在其中再添加一个全局动作。

Fig 9 和 Fig 10(左)展示了不通放大技术处理 Sim1 的放大结果。在这里我们使用一个放大系数 $\alpha = 20$ ，另外，我们在 4.9 - 5 Hz 范围内测试了形变量。在红色矩形框中，我们使用 KLT 方法就可以产生候选动作。在 Fig 9 中，我们展示了蓝色直线的时空片段(见参考帧)。在 Fig 10(左半部分)，SSIM(I_1, I_2) 测量了两幅图像 I_1 和 I_2 之间的结构相似度。这里 SSIM = 1 表明两者的 GT 完全相同，SSIM = 0 表明两者的 GT 完全不同。另外，在 Sim1 中，只有 Fig 9 的黄色矩形之内的区域在估计 SSIM 的时候会被考虑。

Fig 9 表明 DVMAG 最大限度地模仿了 GT，而 Eulerian 和 Youtube 方法产生了模糊误差，Lagrangian 对于动作误差太过敏感。AE 由于缺乏足够长的特征点轨

迹的支持，产生了稳定化不足的结果。

Fig 10(左半部分)表明 Lagrangian 误差和全局动作有相同的轮廓。随着全局动作的增加，该方法的误差也会增加。另外，因为全局动作是最小的，所以这种方法在第 0 帧和第 120 帧的时候达到了最小值。这表明 Lagrangian 放大方法对于动作估计的误差是比较敏感的。Fig 10(左图)中展示的测试结果表明 DVMAG 方法是优于其他的方法的。

Fig 10(右半部分)表明了 Sim1 对于不同的放大系数 α 的不同响应。在这里通过每一个 α 我们都可以预测整个被放大区域中 SSIM 的 GT 平均值。Fig 10(右半部分)表明，相比于起她所有的方法，DVMAG 可以处理更大的放大系数，同时产生更小的误差。比如对于 $\alpha = 20$ 的情况，DVMAG 中的静态误差基本上和 Eulerian 方法，AE 以及 Lagrangian 方法中 $\alpha = 1$ 的误差持平。此外，Fig(10)中的右图表明了 DVMAG 在所有的方法中拥有最慢的下降率。比如说，在 $\alpha = 0 - 40$ 范围内，Youtube, Eulerian 方法和 Lagrangian 方法的跳变比 DVMAG 的更加剧烈。这表明 DVMAG 方法相比于其他所有的方法鲁棒性更好。

Sim2: 我们研究了大幅度动作的放大情况，以及 DVMAG 方法是如何处理这些的。我们测试了 Wu et al 提供的 Guitar 图像序列。这个序列没有任何大幅度动作的图像序列，因为只有琴弦的移动这么微小的动作。我们使用相位方法来放大 Low E note。我们把这种放大结果看作是 GT。跟我们之前将大幅度动作加入到 Sim1 中一样的方法，我们又向 Sim2 中加入了大幅度的全局动作。在这里我们在水平和垂直两个方向上都使用 $A = 50, f = 0.2$ 这两个参数。输出图像序列就是 Sim2。我们分别使用 Eulerian 方法，AE 和 DVMAG 方法来处理 Sim2。我们设 $\alpha = 50$ ，同时我们测试了 72-92Hz 范围内的频率变化。对 DVMAG，我们使用可视流的方法得到候

选动作。在这里整个帧被当作动作掩码。

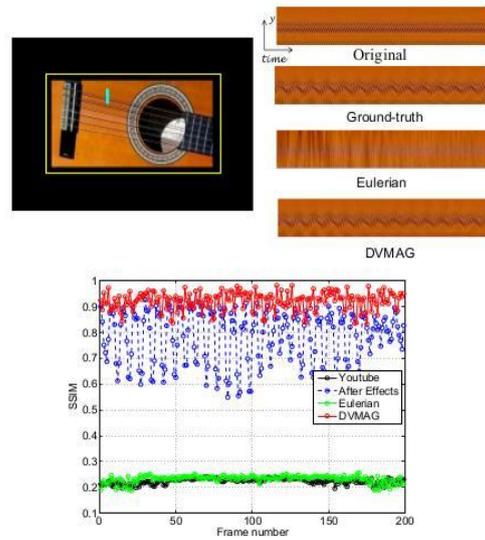


Fig 11(第一行)展示了 Sim2 的一个原始帧及其通过不同方法得到的放大帧。我们展示了蓝色线条的时空片段(见左上角)。琴弦的 Low E note 形变在 GT 模式下显得非常明显。Eulerian 方法产生了明显的模糊误差，而且没有放大琴弦的形变。DVMAG 正确地反映了 GT 而且没有产生模糊误差。Fig 11(下部)展示了每一幅放大后的影像帧基于 GT 的 SSIM。在这里我们只考虑黄色矩形框之内的区域(见 Fig 11 的左上角)。AE 稳定化处理由于错误跟踪而造成了剧烈的跳变。在放大过程之后，这种跳变变得更加剧烈。DVMAG 方法由于其他所有的方法。此外，它的效果比 Eulerian 方法的效果提升大约 200%。(见 Fig 11 的底部)。

5.结论

我们提出了一种新的可以处理大幅度动作的影像放大方法。现在的放大技术在处理大幅度动作的时候都会产生较大的误差。我们的方法先使用临时稳定化技术，之后在进行基于层次模型的放大。去透明化技术被用来放大我们感兴趣的区域，同时保留其余区域的元素不变。结果表明在

处理大幅度动作的时候，如果能够进行艺术的状态方法(什么!)，就会产生更大的放大系数，并且降低误差。未来的工作重点在于对待处理对象的多个不同的动作进行处理。

参考文献

- [1] J. Bai, A. Agarwala, M. Agrawala, and R. Ramamoorthi. Selectively de-animating video. *ACM Trans. Graph.*, 31(4):66:1–66:10, 2012. 2
- [2] G. Balakrishnan, F. Durand, and J. Guttag. Detecting pulse from head motions in video. In *CVPR*, pages 3430–3437, 2013. 1, 2
- [3] M. Grundmann, V. Kwatra, and I. Essa. Auto-directed video stabilization with robust 11 optimal camera paths. In *CVPR*, pages 225–232, 2011. 1, 5
- [4] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *IEEE Transactions on PAMI*, 30(2):228–242, 2008. 3, 4
- [5] X. Li. Image recovery via hybrid sparse representations: A deterministic annealing approach. *IEEE Journal of Selected Topics in Signal Processing*, 5(5):953–962, 2011. 2, 4
- [6] C. Liu. Beyond Pixels: Exploring New Representations and Applications for Motion Analysis. PhD thesis, Massachusetts Institute of Technology, 2009. 2, 3, 5, 6, 7, 8
- [7] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson. Motion magnification. *ACM Trans. Graph.*, 24(3):519–526, 2005. 1, 2, 6
- [8] F. Liu, M. Gleicher, H. Jin, and A. Agarwala. Content-preserving warps for 3d video stabilization. *ACM Trans. Graph.*, 28(3):44:1–44:9, 2009. 1
- [9] F. Liu, M. Gleicher, J. Wang, H. Jin, and A. Agarwala. Sub-space video stabilization. *ACM Trans. Graph.*, 30(1):4:1–4:10, 2011. 1, 5
- [10] S. Liu, L. Yuan, P. Tan, and J. Sun. Bundled camera paths for video stabilization. *ACM Trans. Graph.*, 32(4):78:1–78:10, 2013. 1
- [11] M. Rubinstein, N. Wadhwa, F. Durand, W. Freeman, H. Eugene, and J. Guttag. Revealing invisible changes in the world. *Science*, 339(6119):518–519, February 2013. 1, 2
- [12] J. Shi and C. Tomasi. Good features to track. In *CVPR*, pages 593–600, 1994. 2, 3, 7
- [13] E. Simoncelli and W. Freeman. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In *International Conference on Image Processing (ICIP)*, volume 3, pages 444–447, 1995. 2
- [14] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman. Phase-based video motion processing. *ACM Trans. Graph.*, 32(4), 2013. 1, 2, 3, 4, 5, 8
- [15] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman. Riesz pyramids for fast phase-based video magnification. In *International Conference on Computational Photography*, 2014. 1, 2

- [16] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. 7
- [17] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.*, 31(4):65:1–65:8, 2012. 1, 2, 3, 5, 8