

Simultaneous Optical Flow and Intensity Estimation from an Event Camera

一、简介

事件相机的同步光流和强度估计。提出第一个算法同时恢复运动场和亮度图像，而相机经过一个通用的运动通过任何场景。以前的 DVS 和相关的异步基于时间的图像传感器（atis）都没有针对这种通用估计，而是专注于简化问题。

1.1 采用分布式交换机

用于跟踪简单物体，通常是使用模板拟合从静态相机中或者通过查找当前事件活动的平均像素坐标来解决。但是这些方法不能扩展到跟踪多个、可变和高度纹理图像的独立运动。

1.2 采用 CMOS 相机

使用来自 CMOS 相机的图像为事件跟踪提供目标，以产生视觉里程计系统，这种方法没有从事件数据中进行强度重建。

1.3 光流估计

作者在没有明确估计图像强度的情况下恢复了一个运动场，假设在空间和时间上彼此接近的事件可以相互对应，并局部的将时空平面拟合到这些事件上。在具有锐利边缘和单色块的场景中效果很好，但是在复杂的环境中会遇到麻烦。

1.4 本文方法

文中算法可表述为滑动窗口变分优化。预先定义一个优化时间窗口 T ，然后在此范围内进行精细的时间离散化 δt ，将 T 中所有事件作为输入并联合求解相关时空体中所有单元的速度场 \mathbf{u} 和对数强度 L 。然后将优化向前滑动到高度重叠的位置，将所有单元格的值初始化为先前的估计或预测，并再次求解。

二、具体方法

2.1 估计准备

在亮度恒定下在事件序列的持续时间内估计连续变化的图像速度 \mathbf{u} 和对数强度 L ，其中 $L = \log(I + b)$ ， $I(\mathbf{x} + \delta_t \mathbf{u}, t + \delta_t) = I(\mathbf{x}, t)$ 。

其中， b 为正偏移函数， \mathbf{u} 是像素/秒的单位。

由于在每个像素点的基础上这个方程是不确定的需要引入正则化并执行全局优化以便在整个图像域上实现定义良好的解决方案。

2.2 变分公式

添加正则器本质上是平滑先验，变分方法非常成功的包括平滑先验，平滑先

验使我们能够估计事件之间的图像区域。由于事件与过去的强度有关，即之前的事件与当前事件时间戳的强度有关，因此我们选择将事件测量数据同化为时间窗口内的时空平滑性和光度一致性(光流约束)。由于这两个量是由光流约束耦合的，我们需要联合估计这两个量。我们假设像素处的强度变化仅由光流引起。

$$\begin{aligned} \min_{\mathbf{u}, L} \int_{\Omega} \int_T & \left(\lambda_1 \|\mathbf{u}_{\mathbf{x}}\|_1 + \lambda_2 \|\mathbf{u}_t\|_1 + \lambda_3 \|L_{\mathbf{x}}\|_1 + \right. \\ & \left. \lambda_4 \|\langle L_{\mathbf{x}}, \delta_t \mathbf{u} \rangle + L_t\|_1 + \lambda_5 h_{\theta}(L - L(t_p)) \right) dt d\mathbf{x} \\ & + \int_{\Omega} \sum_{i=2}^{|P(\mathbf{x})|} \|L(t_i) - L(t_{i-1}) - \theta \rho_i\|_1 d\mathbf{x}, \end{aligned} \quad (1)$$

$$h_{\theta}(x) = \begin{cases} |x| - \theta, & \text{if } |x| > \theta \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

$P(\mathbf{x})$ 是在 \mathbf{x} 点触发的所有事件的集合， t_i 和 ρ_i 是 $P(\mathbf{x})$ 中第 i 个元素的时间戳和极性。当达到对数强度的阈值 θ 就会触发事件。 $h(\theta)$ 约束两个事件之间的强度，并将 L 与 $L(t_p)$ 的差值作为输入。

2.3 离散化及最小化

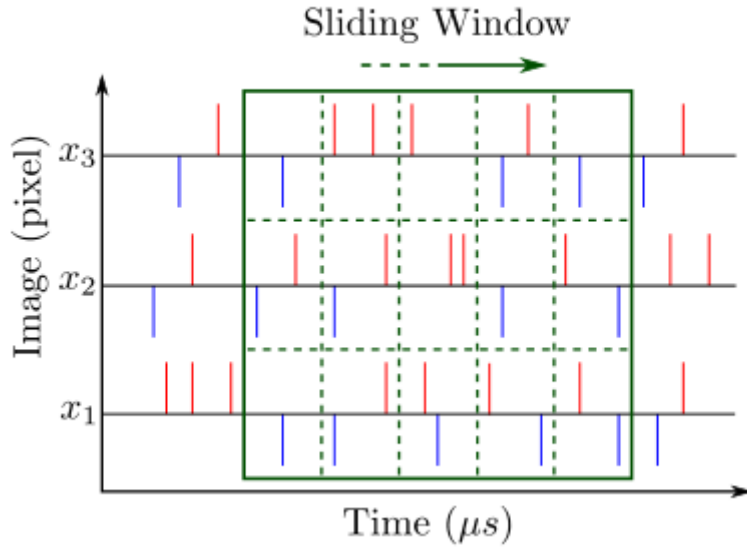


图 1

- 将 Ω 离散为大小为 $M \times N$ 的规则像素网格，将 T 离散为 K 个单元，每个单元的长度为 δt 微秒，从而形成一个时空体积。对于创建的体积中的每个元素，我们通过最小化(公式 1)的离散版本估计 L 和运动 \mathbf{u} 。
- 滑动窗口(绿色框)将传入的积极事件(红色条)和消极事件(蓝色条)放入规则网格(虚线)中。当最小化收敛时，窗口向右移动。

线性插值

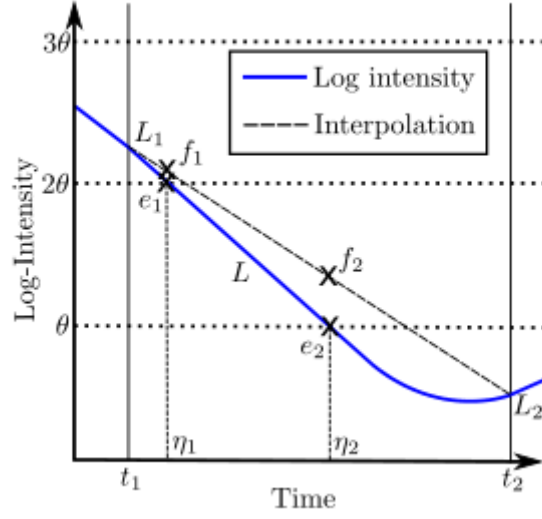


图 2

- 为了使事件数据适应较低的时间分辨率，我们对每个事件发生时的强度进行了线性插值。
- 两个给定事件 e_1 和 e_2 在两个强度估计 L_1 和 L_2 之间的强度近似值。对于离散数据项，我们在每个事件 η_1 和 η_2 发生时使用线性近似 f_1 和 f_2 。

最小化

- 使用预设的原始对偶算法来最小化公式(1)，公式如下：

$$\begin{aligned} \min_{\mathbf{u}, L} \max_{\substack{|\mathbf{a}|_\infty \leq \lambda_1 \\ |\mathbf{b}|_\infty \leq \lambda_2 \\ |\mathbf{c}|_\infty \leq \lambda_3 \\ |\mathbf{d}|_\infty \leq \lambda_4 \\ |\mathbf{y}|_\infty \leq 1}} & \langle \mathbf{D}_x \mathbf{u}, \mathbf{a} \rangle - \delta_{\lambda_1 A}(\mathbf{a}) + \langle \mathbf{D}_t \mathbf{u}, \mathbf{b} \rangle - \\ & \delta_{\lambda_2 B}(\mathbf{b}) + \langle \mathbf{D}'_x L, \mathbf{c} \rangle - \delta_{\lambda_3 C}(\mathbf{c}) + \\ & \langle \mathbf{D}'_x L, \mathbf{u} \delta_t \rangle + \langle \mathbf{D}'_t L, \mathbf{d} \rangle - \delta_{\lambda_4 D}(\mathbf{d}) + \\ & \lambda_6 h_\theta (L - L(t_p)) + \langle \mathbf{E} L - \mathbf{z}, \mathbf{y} \rangle - \delta_Y(\mathbf{y}), \quad (3) \end{aligned}$$

其中， \mathbf{D}_x 和 \mathbf{D}'_x 是关于 \mathbf{X} 的有限差分矩阵， \mathbf{D}_t 和 \mathbf{D}'_t 是关于 t 的差分矩阵。

- 由于 LX 与 \mathbf{u} 的内积，公式（3）中的光流项是双凸的，这里通过交替估计 L 和 \mathbf{u} 来最小化该项，公式（3）最小化方案如下：

$$\left\{ \begin{array}{l} L^{n+1} = (\mathbf{I} + \mathbf{T}_1 \lambda_6 \partial h_\theta)^{-1} (\bar{L}^n - \mathbf{T}_1 (\mathbf{D}'_x{}^\top \mathbf{c}^n \\ \quad - \mathbf{D}'_x{}^\top (\bar{\mathbf{u}}^n \mathbf{d}^n) - \mathbf{D}'_t{}^\top \mathbf{d}^n - \mathbf{E}^\top \mathbf{y}^n)) \\ \bar{L}^{n+1} = 2L^{n+1} - \bar{L}^n \\ \mathbf{u}^{n+1} = (\mathbf{I} + \mathbf{T}_2 \lambda_4 \partial G)^{-1} (\bar{\mathbf{u}}^n - \mathbf{T}_2 (\mathbf{D}'_x{}^\top \mathbf{a}^n \\ \quad - \mathbf{D}'_t{}^\top \mathbf{b}^n)) \\ \bar{\mathbf{u}}^{n+1} = 2\mathbf{u}^{n+1} - \bar{\mathbf{u}}^n \\ \mathbf{a}^{n+1} = (\mathbf{I} + \Sigma_1 \partial F_1^*)^{-1} (\mathbf{a}^n + \Sigma_1 \mathbf{D}_x \bar{\mathbf{u}}^{n+1}) \\ \mathbf{b}^{n+1} = (\mathbf{I} + \Sigma_2 \partial F_2^*)^{-1} (\mathbf{b}^n + \Sigma_2 \mathbf{D}_t \bar{\mathbf{u}}^{n+1}) \\ \mathbf{c}^{n+1} = (\mathbf{I} + \Sigma_3 \partial F_3^*)^{-1} (\mathbf{c}^n + \Sigma_3 \mathbf{D}'_x \bar{L}^{n+1}) \\ \mathbf{d}^{n+1} = (\mathbf{I} + \Sigma_4 \partial F_4^*)^{-1} (\mathbf{d}^n + \Sigma_4 (\langle \mathbf{D}'_x \bar{L}^{n+1}, \bar{\mathbf{u}}^{n+1} \rangle \\ \quad + \mathbf{D}'_t \bar{L}^{n+1})) \\ \mathbf{y}^{n+1} = (\mathbf{I} + \Sigma_5 \partial F_5^*) (\mathbf{d}^n + \Sigma_5 (\mathbf{E} \bar{L}^{n+1} - \mathbf{z})), \end{array} \right. \quad (4)$$

其中， Σ_i 和 \mathbf{T}_i 为对角预处理矩阵， \mathbf{I} 为单位矩阵。 F_i^* 为指示函数， G 为光流项。

三、实验

3.1 参数设置

将滑动窗口 K 设置为 128， δ_t 为 15ms， $\theta = 0.22$ ， $\lambda_1 = 0.02$ ， $\lambda_2 = 0.05$ ， $\lambda_3 = 0.02$ ， $\lambda_4 = 0.2$ ， $\lambda_5 = 0.1$ ， $\lambda_6 = 1.0$ 。所有序列的初始化都假定没有初始运动，只有均匀的灰度强度分布，其中也包括先验图像。将 DVS128 安装在一个标准的基于帧的相机旁边，该相机具有标准的 640×480 分辨率，30 fps 和所有序列的全局快门设置。

3.2 同步估计的好处

比较了在使用公式 (1) 中定义的光流项和不使用时的强度估计，如图 3：

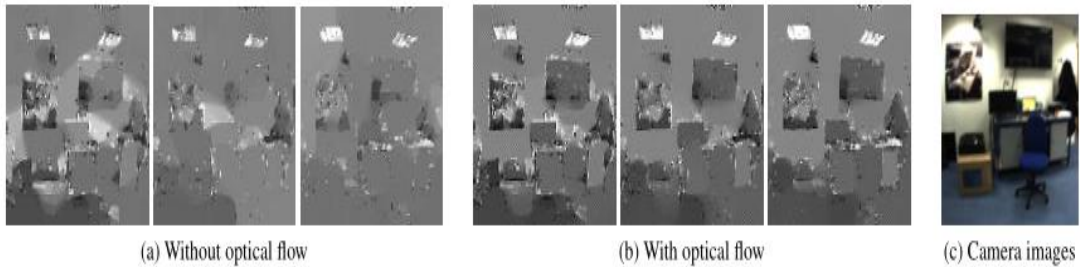


图 3

a. 可以看到，本文方法在没有光流的情况下，仍然可以很好地估计梯度

强的区域的强度，但是在这些区域之间会出现与场景中真实强度不对应的伪影。

3.3 脸序列

在这个序列中展示了强度和速度侦察当相机运动时构建一张移动的脸。如图 4：

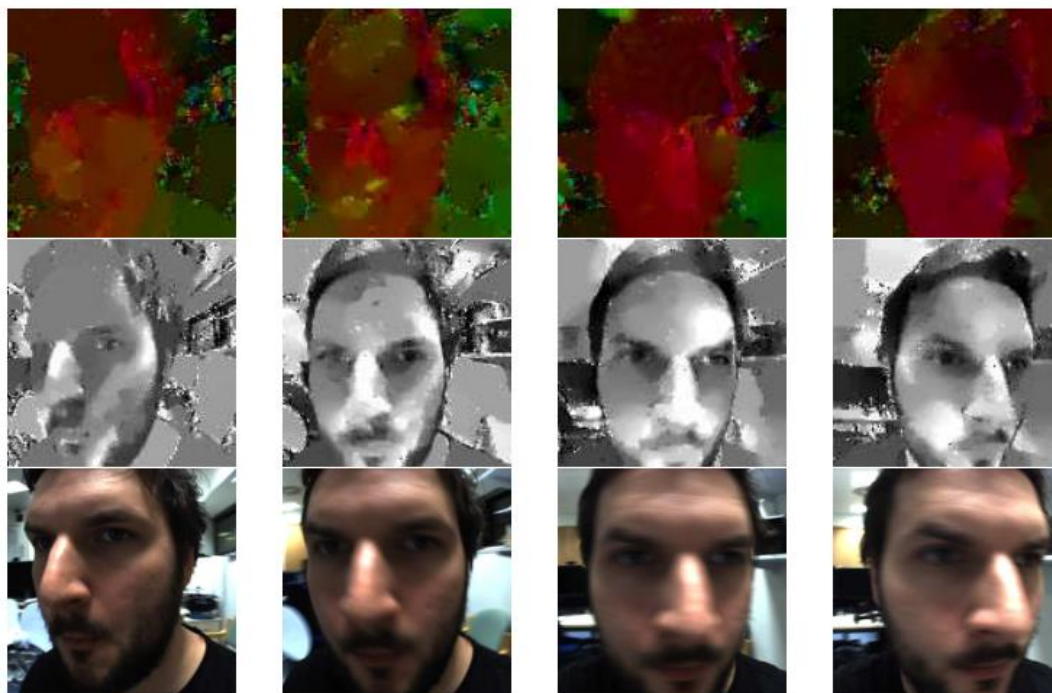


图 4

- a. 可以看到随着处理的事件越来越多在接下来的帧中可以看到更多的细节，速度可视化显示了头部和背景之间清晰的运动边界，证明该方法可以处理运动不连续。
- b. 光强和光流都显示噪声，可能是由异常事件和/或缺失数据引起的。

3.4 高动态范围场景

将本文方法和传统相机在高动态范围场景中进行了比较，如图 5：

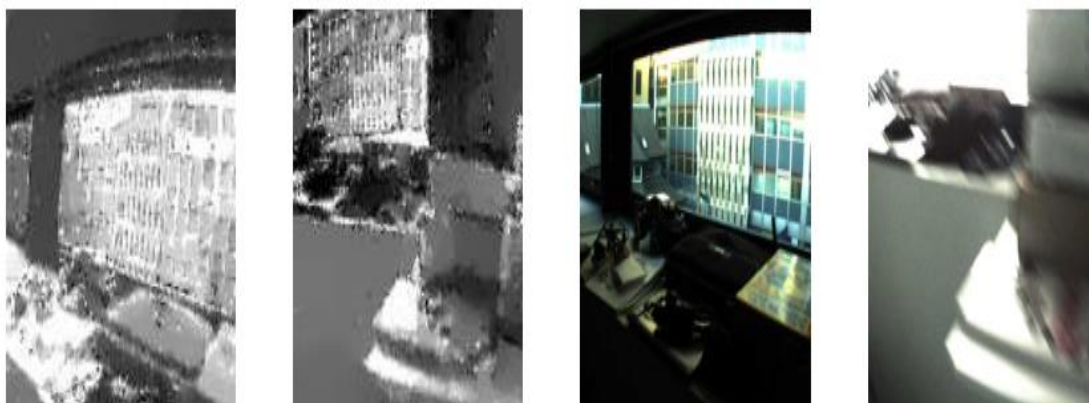


图 5

- a. 将相机从昏暗的房间指向窗外，可以看到文中方法同时恢复室内和窗外的细节，而传统相机由于其低动态范围只能一次显示室内或室外。

3.5 快速运动

展示了该方法在混乱背景前估计快速运动的能力。在桌面场景前扔一个球，同时相机也在运动，如图 6。

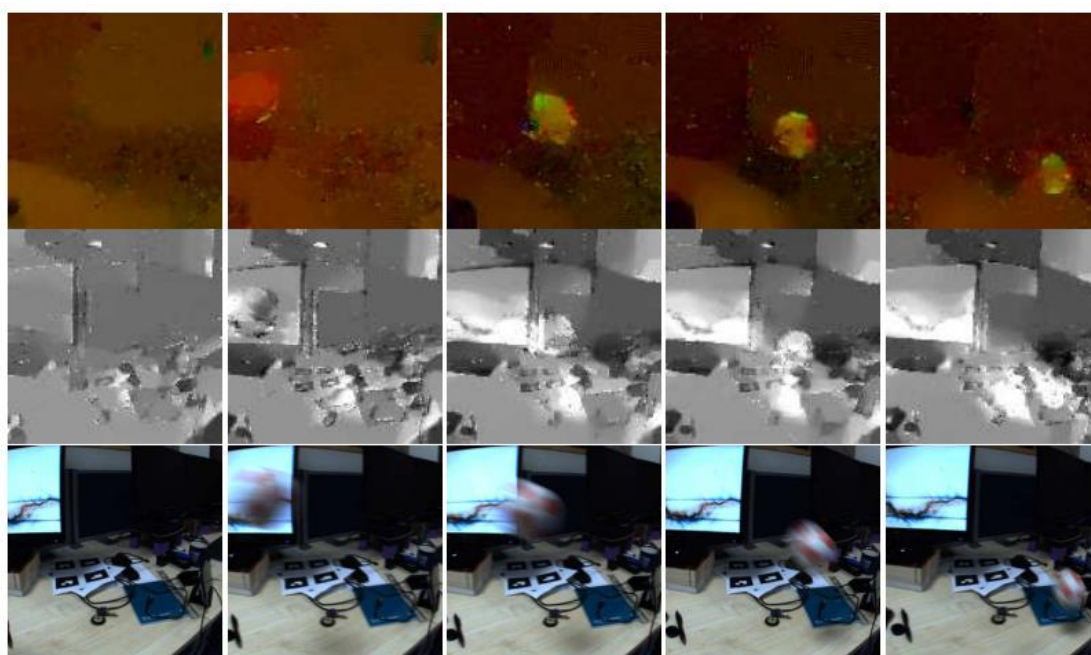


图 6

- a. 在估计的速度场(上)中，我们可以清楚地看到球的分割，而标准摄像机(下)的视频则严重模糊。在强度重建(中)中，我们观察到被球穿过的监视器重建良好，造成了许多事件。
- b. 可以看到传统相机如何受到运动模糊的影响，而该方法能够恢复清晰的运动边界。

3.6 全身运动

展示了一个人在表演开合跳。对于这个序列，将 δ_t 设置为 7ms，并将 λ_1 和 λ_3 降低到 0.01，以避免在光流和强度估计中被平滑。如图 7：

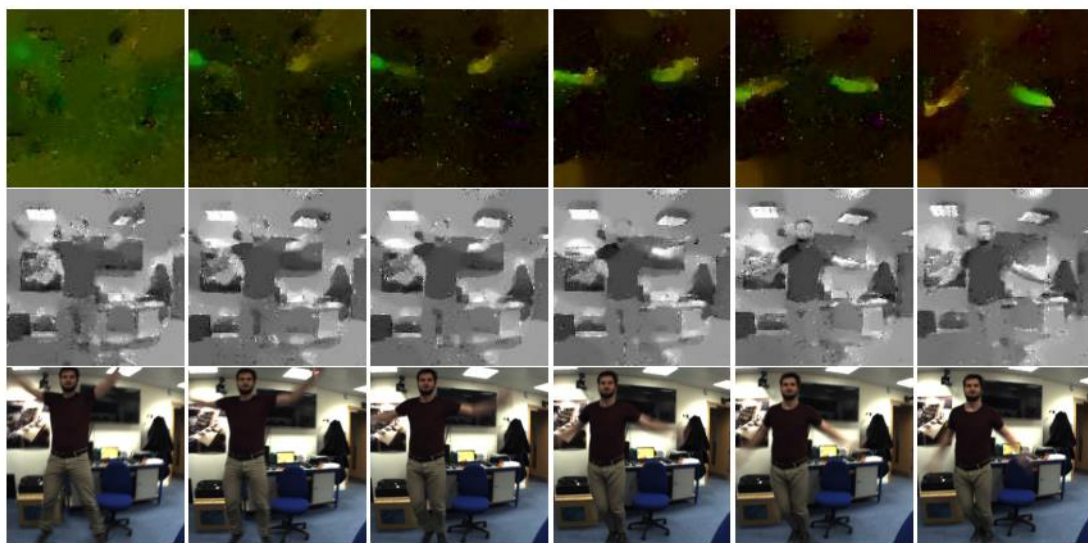


图 7

- a. 可以看到该方法可以很好地估计手臂运动，尽管它占用了很小的图像区域。
- b. 随着平滑权值的减小，噪声事件的影响更强，在运动场中变得明显。

四、总结

证明了来自标准 DVS128 传感器的事件数据可以与密集光流场联合用于重建高动态范围强度帧。基于滑动窗口优化的方法不会对摄像机运动和场景内容施加任何限制。展示了极端，快速运动和高动态范围场景的跟踪和重建，超出了基于帧的相机的能力。因此认为这项工作对于支持事件摄像机将在实时几何视觉中发挥重要作用的说法非常重要——低比特率事件流中的信息确实包含了连续视频的所有内容，甚至更多。

未来方向：研究针对特定问题的特定有效估计算法，如基于模型的 3D 跟踪，3D 重建和 SLAM。