

指导教师： 杨涛

提交时间： 2016/3/21

CVPR2015 Paper Translation

No: 01

姓名： 卢卫卫

学号： 2013302602

班号： 1001136

高效的稀疏到密集光流估计使用学到的基础和图层

乔纳斯武尔夫 迈克尔·黑 马克斯普朗克研究所的智能系统，德国图宾根
{jonas.wulff, black}@tue.mpg.de

摘要

我们专注通过采用一种稀疏到密集地方法使得评价光流这一难以实现的目标更精确高效。给定一组稀疏的比赛，我们回归使用一套已了解的全框架的基础上流动领域。我们从四家好莱坞电影公司学习到通过流计算了解自然界流动领域的主要部分的方法。光流场是简洁的近似基础流动领域的加权和。我们的新的 PCA-Flow 算法从稀疏特征匹配稳健估计这些权重。这种方法运行在基于单 CPU 使用 MPI-Sintel 数据集 200ms 的框架上更精确，并且更加显著的比流行的 LDOF 和 Classic+NL 算法速度快。然而，对于某些应用程序来说，结果是太光滑了。因此，我们制定了一种新的稀疏分层方法，每一层都是由其主流成分代表。不同于现有的分层方法，估计是快速的因为它只使用稀疏匹配。我们将不同层信息合并到使用图像意识的流变密集流领域。由此产生的 PCA 层方法在 3,2s/框架运行，比 PCA-Flow 更精确，并且获得在闭塞地区的状态艺术表现。

1. 简介

光流场的最新进展已导致增加准确性，驱动部分的基准，如

Middlebury[3]，MPI-Sintel[10]和 KITTI[16]。特别是，最近的方法使用稀疏或稠密匹配捕获远程运动同时利用传统的变分技术，从而得到高精度 [9、24、26、28、29、50、53]。还有其他使用分层模型或分割区域方法进行关于咬合关系的理由，并更好地估计议案在边界和无与伦比的地区 [24、28、43、46]。然而，在许多应用中，速度是至少同样重要。最准确的方法需要几秒到几分钟每帧。规避方法往往不准确或需要 GPU（或两者）。

为了解决精度和速度，我们建议一种新的稀疏到密集方法，基于稀疏特征匹配紧跟插值。稀疏特征是规避鲁棒计算，可以捕获远程运动。通过插值之间这些稀疏的火柴，密集流可以计算上有效。然而，由于稀疏匹配和不均匀覆盖的图像中的异常数据，泛型抽取做不好工作。相反，我们学会从培训光流场前瞻通过主成分分析 (PCA) 插值。

学习线性模型的流场的想法并不是新 [7, 15]，但以前的工作应用这种模式仅在图像修补程序不完整映像。为了训练我们的 PCA 模型我们使用光流场计算从 8 个小时的视频帧从使用现有的流场算法 (GPUflow [52]) 的四个商业电影。要培训流噪声处理

我们好给我们巨大的训练集 [21] 使用鳞片鲁棒 PCA 方法。

我们的方法计算密集流通过估计的位置在 PCA 子空间最能解释稀疏匹配 (图 1(c))。在第一批, 它并不立即明显那个可以表示泛型流场经久使用从计算流场; 构造成为 PCA 基础我们证明这确实工作。

这种做法是非常规避。我们新型的流场算法, 称为 PCA-流, 已运行约 190 毫秒每个框架上标准的 CPU; 这是最快的基于 CPU 的方法上叻 [16] 和 MPI-辛 [10]。那里是权衡精度和速度, 而 PCAFlow 比不上最准确的方法, 它是有力地比上叻下一个的最快方法更准确, 是比最近, 广泛使用的方法, 如 LDOF [9] 和经典 + NL [44] 在 MPI 辛更准确。有趣的是, 通过学习足够的数, 我们获得较小的错误, 比用来训练我们的主成分分析基础的算法。

快速、准确, 可用于许多任务的 sufficiently 同时 PCA 流动不包含高频率空间信息, 因此过度平滑运动边界流场。若要同时保留工件获得大幅运动边界, 我们建议每个图层在哪里估计从稀疏匹配的一个子集的 PCA 流领域的新型分层流场模型。以前的分层的模型是计算昂贵 [45, 46]。通过使用稀疏匹配和博学的 PCA 插, 每一层可以动议上有效计算使用期望最大化 (EM) [23]。

为了计算单项密集流领域, 我们必须结合流场前瞻估计为每个图层。

我们使用马尔科夫随机 (MRF) 包含了图像证据中每个像素的 PCA 流经久选择这样做。此层 PCA 方法计算光流场前瞻与很多尖锐的运动边界并减少了总体的错误 (图 1(d))。同时, 它仍然是相当快, 以 MPI 辛平均板/帧。在辛上它是比像 MDP-Flow2 [53], [4] 简明, MLDP 最近方法更准确-[33] 的经典 + NLP [44] 与传统分层方法, FC-2Layers-FF [46], 这是至少两个数量级速度较慢。最有趣的是, PCA 层是特别好 (无与伦比) 闭塞地区实现较低的错误比 DeepFlow [50] 对辛。

为研究目的, 方法和博学的流场基础的代码索取 [1]。

2. 以前的工作

光流场和稀疏特征跟踪有悠久的历史。在这里我们关注的元素和组合最关乎我们的方法。传统的光流场的变分方法取得好的成绩, 顺利和小的运动, 但在远距离运动失败。特征匹配, 另一方面, 需要 sufficient 图像局部结构, 并因此只收益率稀疏的结果。因此, 他们必须有“densified”, 通过显式插值/回归分析, 或通过在一个变框架集成。

在光流场的稀疏特征。使用跟踪的功能来估计运动的想法的根源在早期信号处理 [5]。早期的光流场方法用相关匹配处理大运动 [2]。

吉布森和 Spann [18] 描述了一种两阶段方法, 第一批估计稀疏特征轨

迹，随之而来的插值阶段。他们跟踪阶段使用 MRF 强制时空平滑，而插值阶段基本上是优化传统密集的光流场的一个目标。这使得该方法计算代价昂贵。

Nicolescu 和梅迪奥尼 [34] 使用特征匹配得到候选运动并使用张量投票为插值。他们然后分割成仅使用的流场信息的区域的流场。我们 PCA 流方法还有类似的阶段，但是用于 densification 的博学的 PCA 基础。

郎等人 [26] 第一批图像序列跟踪稀疏特征，然后使用图像引导的滤波方法功能之间进行插值。他们依靠世俗平整度在多个帧。其结果是视觉上吸引人，但他们明德报告性能差，不做评价辛或叻。

Leordeanu 等人 [28] 在网格上使用 k-近邻到夏装对应的特征和迭代 refine 估计的本地 affine 运动和闭塞。他们遵循这标准变分的光流场法 [44]。他们的算法需要每个帧上辛对 39 分钟。

几种方法结合特征匹配和传统的变分方法，在单一的优化。刘等人 [29] 稠密 SIFT 特征结合光学 flowbased 正则化。在这两个框架使用区域和猪描述符分割的一下和马利克 [9] 比赛区域。这些描述符匹配，然后在他们密集流场优化形成额外数据术语。肯尼迪和泰勒 [24] 使用传统数据术语在三角化补丁和稠密的猪匹配；他们的方法，TF + OFM 有着良好的 MPI 辛但计算代价昂贵（叻 350s）。

Weinzaepfel 等人 [50] 使用类似的方法，但提出了一种新匹配机制被称为 DeepFlow。徐等人 [53] 使用稀疏的 SIFT 特征生成额外的平移流假设。他们然后使用标签的方法来将像素分配给那些假设之一或更传统的变分流量。

基于 PatchMatch 的方法介于密集的光流场和稀疏特征估计 [11, 30]。他们计算密集但近似函授 \mathbf{f}_i 高龄和 refine 使用各向异性扩散 [30] 或分割成小表面修补程序 [11] 这个领域。

快速计算流场。扎克等人 [54] 进行的第一批来证明在 GPU 上的光流场计算。他们显示如何总变分方法可以并行与着色器使用传统的目标函数。他们取得小决议（ 320×240 像素）的实时性能。Werlberger 等人 [52] 该方法扩展到鲁棒经常化。在最近的 GPU，他们的算法需要大约 2 秒每帧，分辨率为 1024×436 像素。Rannacher [35] 提出了一种非常快速的方法，但需要立体图像和预计算的差距领域（例如，使用 FPGA）。孙达拉姆等人 [47] 到 GPU，与运行时的 1.8 秒 640×480 像素的图像对港大位移光流场方法 [9]。报告运行时依赖于图像的分辨率和 GPU 的类型。

陶等人 [48] 算法与图像输入分辨率 sublinearly 鳞片，通过计算图像像素的所选子集上流量和插值的剩余的像素。然而，它仍有 2 秒左右的

运行时间在明德图像对。宝等人 [4] 使用 GPU 进行关于 0.25s 年/帧在辛上运行在他们最近的简明方法。我们基本的方法是稍微不那么准确，但约 60ms 更快，并且不需要 GPU。我们分层的方法更精确，但需要板/帧。

非本地平滑和图层。光流场正则化通常使用小街区，但最近的工作表明，非局部正则化的价值。这可以通过中值滤波 [44, 51] 或密被连接的磁流变液 [25]。在这里我们实现非局部平滑使用学到的主成分分析基础。

分层的模型 [13、23、42、45、46、49] 提供另一种方法。分层模型的优点是分割涉及现场几何。然而，当前方法的缺点是运行时之间数十分钟 [46] 变化几十个小时 [45]。

学习空间模型的流场。简单的线性模型（平移或 affine）的光流场图像修补程序中有着悠久的历史 [31]。随着斑块面积的增长，复杂的运动，也是如此，这种模式不再适当的 [23]。因此这种线性模型通常只能在本地使用。

舰队等人 [15] 通过使用 PCA 学习基础流场前瞻从例子扩展线性建模方法。他们估计这些修补程序使用计算代价昂贵的翘曲-基于优化方案中的模型的系数。我们使用稀疏特征的方法是更多的规避和可以应付长范围对应。Chessa 等人 [12] 使用基础流场前瞻中局部贴片，占 affine 运动再加上现场的几何结构变形。几位作者

探索了 PCA 模型的行走人类 [14, 20]。

罗伯茨等人 [37] 学习光流场为 egomotion 使用概率 PCA 子空间。他们使用这个子空间，估计从稀疏的运动估计的密集流领域。他们限制自己对 egomotion、火车和试验相似的序列捕获从同一个平台，并使用只有一个二维的子空间的 45×13 像素分辨率很低。最近的工作延伸，但侧重于语义来成障碍类从这议案子空间，而不是精确的运动估计 [36]。

除了这些线性模型，还有小工作学习空间模型的全流场。罗斯和黑色 [39]，空间定期学习高阶 MRF 模型光流场，但社区化是仍然很小（ 5×5 个像素）。罗森鲍姆等人 [38] 学习空间模型的流场和这用于去噪而不估计的流量。吉布森和马克斯 [19] 使用 overcomplete 字典的光流场在修补程序正规化流场。

3. 光流场学的基础

我们的基本假设是，光流场前瞻可以近似地认为加权数相对较小的基础流场前瞻 b_n , $n = 1 \dots N$ ，与相应的权重 w_n

$$\mathbf{u} \approx \sum_{n=1}^N w_n \mathbf{b}_n.$$

在这里， \mathbf{u} 和 \mathbf{b}_n 矢量光流领域，包含水平和垂直运动的叠列向量： $\mathbf{U} = \mathbf{U}^T$ \mathbf{x} , $\mathbf{u}^T \mathbf{Y}^T$ 。我们假设的水平和垂直流

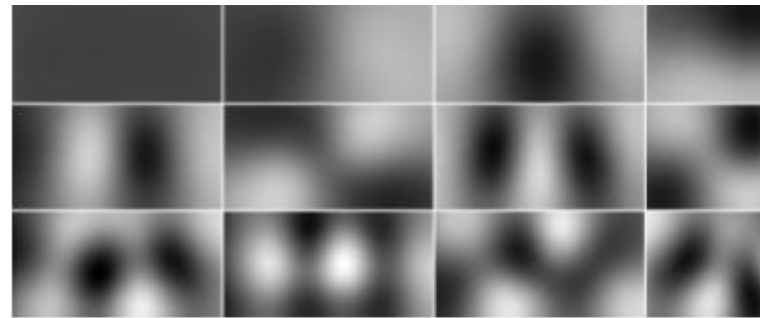
分离元件的基向量，使水平运动有 $\{e_1, \dots, e_{n-1}\}$ ，和垂直的 $\{e_n, e_{n+1}, \dots, e_{2n}\}$

3.1 学习的流量基础

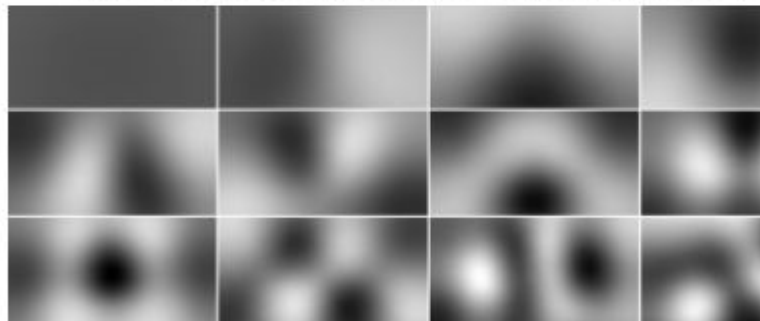
学习基础流场前瞻，我们使用来自四个好莱坞电影跨越几个流派（《星球大战》，《巴贝尔》，《不朽的园丁》，《鱼米之乡》）的数据。对于每一部电影，我们计算光流场使用 GPUFlow [52]。此方法不是最精确的（正如我们将看到，它是比我们用它训练的 PCA 流算法不太准确）。然而，它是最快的方法，与可用的参考实现，并且有大约 2 秒每帧运行。计算光流场需要大约 4 天每部电影。计算流场然后调整大小为 512×256 像素和流值的大小进行缩放相应；这是在 [50] 所使用的相同的分辨率。

从计算光流场前瞻，我们随机选择 180,000 帧，受到最大供我们使用的内存量的限制。我们第一批减去平均流，其中包含由 GPUFlow 方法引起一些一致边界伪影。请注意，在这里，我们的组件的维数大于数据点的数量。然而，相比理论维度，我们提取只有非常少量的主成分，这里 $N = 500$ 、250 的水平运动和垂直 250。由于计算光流场包含异常值由于编辑削减及机架的光流场计算的失败，我们使用一种鲁棒的 PCA 方法来计算主成分 [21]。提取 500 个组件所需的总时间是大约 22 个小时。然而，这已经做唯一的一次，offline；我们做学的基础可用 [1]。图 2 显示了第

一批 12 流组件在水平和垂直方向。请注意，其中一个还可以培养一个



(a) Principal components for horizontal motion



(b) Principal components for vertical motion

图 2：第一 12 组件为水平和垂直的运动。对比增强的可视化



(a) Ground truth optical flow (b) Projected optical flow

图 3：示例突出地面真理流拖到第一批 500 主要组件上。

纵向和横向运动的联合的基础。然而，在我们的实验中，单独的基地一直高于联合的基础。此外注意到的第六个组件不直接对应于 affine 议案，与小流量修补程序 [15] 发现了什么。

图 2 显示生成的主成分类似于基函数的离散余弦变换 (DCT)。为了向我们学的基础，与相同数量的组件实现

相媲美的性能，我们生成一个 DCT 依据与十倍更多的组件和基础追求用于选择最有用的。尽管这样，DCT 基础给稍差的终结点错误在我们的实验，所以我们并不认为它进一步。

图 3 显示地面真实流场领域的投影从辛的教训的基础上。请注意，没有辛训练数据用的学习基础，所以这个测试泛化。此外注意到辛序列是相当复杂和预计的流场是得更顺畅；这是可以预料。对于数目主成分对重建精度的影响以及与 DCT 基础的定量比较，请参见补充材料 [1]。

4. 估计流场

我们那移位光流场估计系数给出图像序列和博学的流量基础。为此，我们首先计算稀疏特征匹配建立对应的两个帧之间的关键点。然后，我们估计产生两个匹配的场景运动和光流场经久的一般结构相一致的密集流领域的系数。

4.1. 稀疏特征匹配

我们的算法开始通过估计 K 稀疏特征匹配跨相邻帧；即对点 $\{(p_k, q_k)\}$, $k = 1 \dots K$. p_k 是帧 1, (通常以可视方式是不同的) 特征点的二维位置和 q_k 是相应的特征点定位在框架 2 中。每个这些书信诱导位移向量 $v_k = q_k - p_k = (y, x, v_k)$ 使用稀疏特性有两个主要优点。首先，它提供图像和对象运动的粗的估计而计算是相对便宜。第二，它建立了长时间的范围对应，是传统的、致

密的流场的方法，估计 [9、50、53] 陈设。首先，我们规范使用 CLAHE [55] 增加可以通过特征探测器捕捉到的细节图像的对比度。然后，我们使用的功能从 [17]，这专为视觉里程计应用。我们发现，若要跨视频帧匹配功能，这些功能工作比图像匹配功能，如冲浪或 SIFT 好得多。后者是不变的对种类繁多的几何变形，很少发生在相邻帧的视频，并因此返回大量的不匹配。此外，我们使用的特点是计算要便宜得多：目前，匹配不同的两个帧在 MPI 辛平均 80 女士使用稀疏特性创建覆盖率低的问题在非结构化的图像区域。然而，在古典的光流场中也存在此问题：如果是缺掉的图像结构、数据期限变得含糊不清，和流场计算依赖正则化，就像我们的方法依赖于插值的学术基础。特征匹配将始终包含异常值。我们解释这些匹配过程中通过制定我国优化鲁棒地下面。图 4 (a) 显示的帧和图 4 (c) 及相应的功能。蓝色所示的功能有误差小于 3 个像素；功能有更大的错误是红色的。

4.2. 回归：由疏到密

扩展到密集的光流场领域稀疏特征匹配是回归问题。使用我们学的流基础矢量 b_n ，这可以制定作为加权线性组合的流量基础上向量，最好

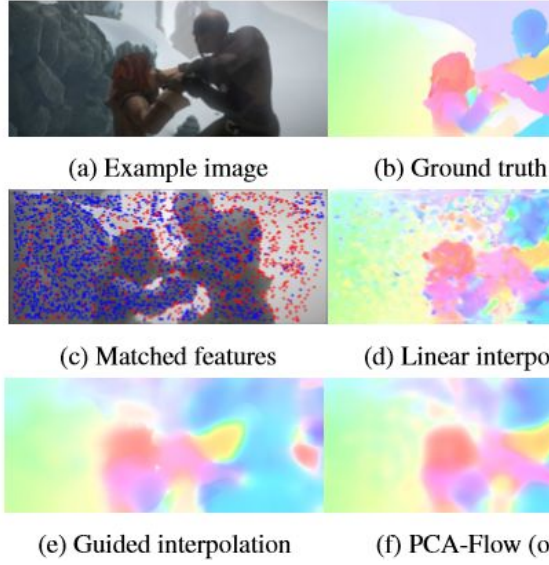


图 4 : 稀疏特征和可能篡改解释了检测到的特征匹配。重量然后移位密集流场。首先, 考虑基本版本的方法。这可以表示为一个简单的最小二乘问题在未知 $w = (w_1, \dots, w_N)$:

$$\hat{w} = \underset{w}{\operatorname{argmin}} \|Aw - y\|_2^2 \quad (2)$$

$$A = \begin{pmatrix} \mathbf{b}_{1,x}(\mathbf{p}_1) & & \mathbf{b}_{N,x}(\mathbf{p}_1) \\ \vdots & & \vdots \\ \mathbf{b}_{1,x}(\mathbf{p}_K) & \cdots & \mathbf{b}_{N,x}(\mathbf{p}_K) \\ \mathbf{b}_{1,y}(\mathbf{p}_1) & & \mathbf{b}_{N,y}(\mathbf{p}_1) \\ \vdots & & \vdots \\ \mathbf{b}_{1,y}(\mathbf{p}_K) & & \mathbf{b}_{N,y}(\mathbf{p}_K) \end{pmatrix}. \quad (3)$$

和

我们使用最近邻点法来计算元素在分数的坐标; 更好的插值并没有提高我们实验的准确性。 $y = (v_1, x, \dots, v_K, x, v_1, y, \dots, v_K, y)^\top$ 包含匹配点的运动。

\hat{w} , 进而估计密集的光流场求解方程 (2) 收益率

$$\mathbf{u}_{est} = \sum_{n=1}^N \hat{w}_n \mathbf{b}_n. \quad (4)$$

不幸的是, 稀疏特征匹配通常包含异常值。因为搜索功能火柴都是跨整个

图像 (即功能运动空间范围不限制), 则场糟糕的比赛引起的误差往往规模庞大, 因此可以有香港的情商 (2) 的解。因此, 我们解决了一个鲁棒的版本的情商 (2)

$$\hat{w} = \underset{w}{\operatorname{argmin}} \rho(\|Aw - y\|_2^2) \quad (5)$$

其中 $\rho(\cdot)$ 是鲁棒的柯西函数

$$\rho(x^2) = \frac{\sigma^2}{2} \log \left[1 + \left(\frac{x}{\sigma} \right)^2 \right]. \quad (6)$$

参数 σ 控制对异常值的敏感性。请注意, (6) 是许多可能的鲁棒估计 [6] 之一。我们发现柯西估计很好地工作。如果输入的图像有不同的分辨率比流场基础上, 我们第一批检测全分辨率, 功能和规模及其位置我们流基础的决议。重量然后估计在此分辨率下, 和由此产生的光流场田野是采样和再次缩放到输入图像的分辨率。请注意, 人也可以估计系数使用经典的密集估计的参数化光流场 [31]。这是 [15] 中使用的方法。我们实现了此方法, 发现, 奇怪的是, 其准确性是相当于我们 PCA 流方法与稀疏特征匹配。因为它是要慢得多, 我们不认为它进一步在这里; 看到结果和比较 [1] 的补充材料。

4.3. 实行事先

方程 (5) 不会考虑到 w 的分布。对 w 简单事先给出特征值计算期间 PCA 对训练流场前瞻。新的序列可能有相当不同的统计信息, 但是。在叻, 例如, 运动纯粹由一辆车, egomotion 引起的并因此少一般比我们训练数据。

虽然叻和 MPI 辛包含培训数据，数据量降糖学会充分流基础。然而，我们可以保持基础固定和适应事先。自 500 维子空间定义为按我们流场计算中的事先谎言，这就需要培训数据少得多。地面真实 流场前瞻（例如从叻或辛），我们项目这些我们通用流场的基础上进行了计算 Γ ，逆协方差矩阵的系数。我们表达我们事先对 w ：

$$\hat{w} = \underset{w}{\operatorname{argmin}} \rho(\|Aw - y\|_2^2) + \lambda \|\Gamma w\|_2.$$

直观地说，如果某些系数变化不大投影的训练集流场基础上，我们限制此系数对小值在推理过程中。当训练数据可用时，此正规化提高性能有力地。

我们求解方程 (7) 使用迭代加权最小二乘和引用作为 PCA 流的方法。图 4 显示了我们方法（与两个简单方法，到密流领域从稀疏特征进行插值，最近邻插值 (4(d)) 和图像引导插值 [22] (4(e)) 4(f)) 的结果。这些通用的内插方法无法检测和剔除野值引起错误的特征匹配。因此，他们平均端点误差对辛测试集（线性：9.07 px；引导：8.44 px）高于我们基本的方法，PCA 流 (7.74 px)。

5. 稠密的运动，从稀疏的层数

虽然顺利流场前瞻 PCA 流产生可能适合于某些应用程序，许多应用程序要求精确运动边界。为此，我们制订的方法生成建议使用分层的方法，将它们合并使用图像证据。

5.1. 稀疏图层

在这里，我们假设，全光流场领域由 M 简单的运动，其中的动作之一分配给每个像素组成。在每一层流量由如上与一个 modification 我们学到的基础：因为每一层的运动应该是比简单的全流场，我们改变事先。获得分层的表示的培训，我们第一批群集运动前瞻中我们设置成层具有类似运动的培训。然后，我们为每个层计算 w ，计算各阶层的权值的协方差矩阵 Σ 跨整个训练集，并使用 $\Gamma = \Sigma^{-1}$ 在情商 7。计算简单运动在测试时，我们第一批群集稀疏特征匹配与硬 assignments1 使用 EM 算法。若要初始化，我们进 M 集群使用 K -均值群集功能。分配功能到图层在迭代我被表示为工作分配变量联系 k ， $k = 1, \dots, K$ ，哪里联系 $k = m$ 特征点 p_k 的手段分配到层 m 。一套层特征点 p_k 到迭代中的层 m 的距离我是鉴于作为

$$d^{(i)}(p_k, m) = \|u_m^{(i-1)}(p_k) - v_k\|_2 + \alpha \|p_k - \operatorname{median}(p_k | a_k^{(i-1)} = m)\|_2 \quad (8)$$

嗯是光流场层 m ， v_k 的特征位移作为上述的愈合。右侧部分的距离是，对所有的功能分配到在前一次迭代； m 中位数点 p_k 最初，中位数都已初始化为中心的形象。A 是一个权重因子。特点是然后被分配到图层

$$a_k^{(i)} = \underset{\hat{m}}{\operatorname{argmin}} d^{(i)}(p_k, \hat{m}) \quad (9)$$

和图层都将更新为

$$w_m^{(i)} = \operatorname{estimate}(\{p_k | a_k^{(i)} = m\}) \quad (10)$$

在 $\text{estimate}(\cdot)$ 解决了情商 (7) 使用给定的功能的子集。由于每一层的运动比简单运动的整个图像，图层不需要捕获合身空间细节。因此我们减少线性系数的数量从 500 到 100; 这是 sufficient 为较好的效果。我们以迭代方式解决情商的人。8) - (10) 20 次迭代，或直到分配正义与发展党不要更改了。

5.2. 结合层

估计的层运动给他们分配的功能，但我们创建一个新的插值问题。我们不知道非特征像素对应于哪一层。因此我们开发一种方法来将图层合并到密流领域。在文献中对相关问题 [27, 32] 提出了几种方法。在这里，我们使用一个简单的 MRF 模型。

层估计步骤生成 M 近似的光流场前瞻，由其系数 w_m 和单项赋值变量 a_k ，表示的稀疏特征属于哪个运动模型。我们对待每个图层的运动作为一项提案。除了这些 M 流场前瞻，我们计算两个额外的流量建议：简单的单应矩阵模型、鲁棒通常指所有匹配的特征和近似全流场现场的臆测，即求解方程 (7) 的所有功能和 500 个主要组件 (“全球模式”)。因此，标点符号 $M = M + 2$ 。

在每个图像位置 x ，这项任务是现在分配标签 $l \in 1 \dots M$ ，提供最佳的流场模型，在此像素。然后，作为给出单项光流场领域 $u_{final}(x)$

$$u_{final}(x) = \sum_{m=1}^M \delta[m = l(x)] u_m(x). \quad (11)$$

找到可以作为能量最小化问题制定琐，这可以解决多类图通过削减 [8]:

$$\hat{l} = \underset{l}{\operatorname{argmin}} \sum_x E_u(x, l(x)) + \gamma \sum_{y \in n(x)} E_p(x, y, l(x), l(y)) \quad (12)$$

其中 E_u 和 E_p 分别为一元和成对的能量，而类表示四街坊的十、省略参数 $x, l(x)$ ，unaries 欧盟是作为愈合

$$E_u = E_{warp} + \gamma_c E_{col} + \gamma_l E_{loc}. \quad (13)$$

翘曲的成本。这种扭曲成本 E_{warp} 是一个 rectified 亮度和梯度恒常性术语：

$$E_{warp}(x, l(x)) = 1 - \exp\left(-\left(\frac{c(x, l(x))}{\sigma_w}\right)^2\right) \quad (14)$$

$$c(x, l(x)) = \left\| \begin{pmatrix} I_1(x) - I_2(x + u_l(x)) \\ \nabla_x I_1(x) - \nabla_x I_2(x + u_l(x)) \\ \nabla_y I_1(x) - \nabla_y I_2(x + u_l(x)) \end{pmatrix} \right\|_2 \quad (15)$$

彩色成本。我们建立每个图层，使用分配到该图层的功能点的像素颜色外观模型。这有助于特别是在闭塞地区，翘曲误差高，但颜色提供关于哪一层，使用一个好的线索。

$$E_{col}(x, l(x)) = -\log p_{l(x)}(I_1(x)) \quad (16)$$

$$p_{l(x)}(I_1(x)) = \mathcal{N}(\mu_k, \Sigma_k) \quad (17)$$

Σ_m 哪里 μ_m ，计算从像素 $\{I_1(p_k) | a_k = m\}$ 在这里，我们使用简单的多元高斯分布作为颜色模型; 我们发现这些执行甚至高斯混合模型的多组分更好。对于单应矩阵模型，我们那里的环境内的所有功能; 分布对于全球的模型，我们那里的环境它的所有功能。

功能位置成本。最后，我们注意到分配给给定层的特点经常空间聚集，模型的质量减少，远从功能区域。因此，我们鼓励空间压实度的图层，使用

$$E_{loc}(x, l(x)) =$$

$$1 - \sum_{k|a_k=l(x)} \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp\left(-\frac{(x-p_k)^2}{\sigma_l^2}\right) \quad (18)$$

对于单应矩阵模型，我们再一次使用只内功能，对于全局模型中，我们使用所有。

图像调制的平滑度。若要执行空间平滑，我们使用的图像调制的成对 Potts 能量，从 GrabCut [40]:

$$E_p(x, y, l(x), l(y)) = -\delta[l(x) = l(y)] \exp\left(-\frac{(I_1(x) - I_2(y))^2}{2E[\|\nabla I_1\|_2^2]}\right) \quad (19)$$

与电子 $[\cdot]$ 表示预期值。这种能量鼓励空间平滑的层标签之间的像素，除非有一个强的梯度图像中。因此，它允许层标签来更改图像边界。

6. 评价

本节在标准的光流场数据集上我们在精度的算法的性能。此外，我们提供的运行时信息，并且这涉及其他当前的光流算法。使用上可用的训练集的交叉验证，确定所有参数。对于 PCA 层，我们使用 $M = 6$ 层。数目的主成分特征匹配，以及更多的视觉效果，影响实验的其他参数值，请参阅补充材料 [1]。

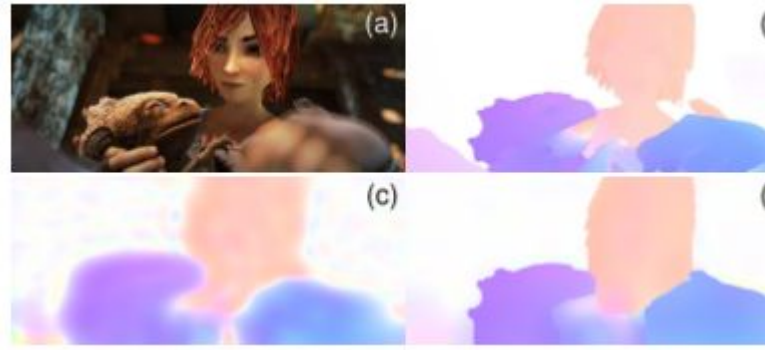


图 5 : 结果对 MPI 辛: (a) 的图像; (b) 地面真实流场; (c) 主成分分析-流程; (d) PCA 层。

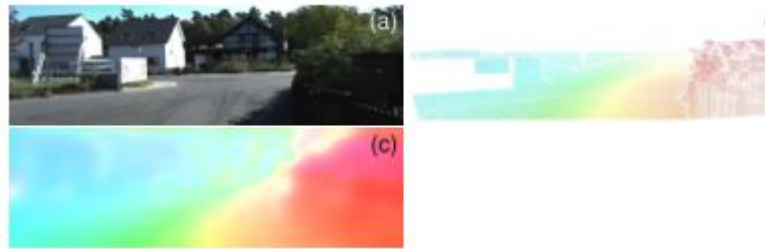


图 6 : 结果叻: (a) 的图像; (b) 地面真实流场; (c) 主成分分析-流程; (d) PCA 层。

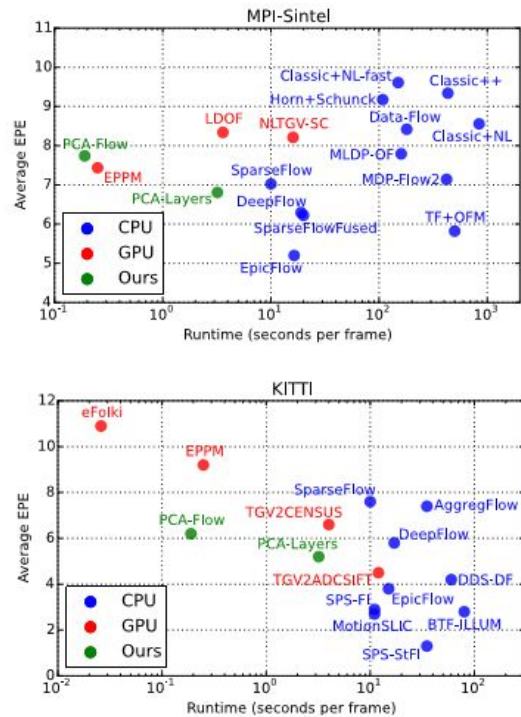


图 7 : 平均 EPE 与运行时上辛和叻

6.1. 在光流场数据集上的评价

MPI-辛。图 5 显示了从清洁通 PCA 流和 PCA 层的 MPI 辛光流场基准的训练集的示例。PCA 流面临过平滑光流场 f_i 高龄，但能正确地估计大部分的长距离运动。通过计算，结合多个图层，PCA 层是能够计算好运动和精确定位运动边界。

上辛测试集，PCA 流目前排名 36 对这两个单项传递地点 22 (EPE = 8.65 px) 和清洁通 (EPE = 6.83 px)。虽然不是最准确的方法，它只需要 190 ms 每帧，而一贯表现优于 LDOF 和经典 + NL 的常用的方法。值得注意的是，我们的表现优于 GPUFlow [54]，我们用来生成训练数据。GPUFlow 以 2 s 每个框架，并达到 12.64 平均 EPE px (干净) 和 11.93 px (单项)。我们的方法生成的光流场田野以来空间分辨率很低，我们把它比作经典 + NLP 在图像分辨率的 64 x 32 像素。在此分辨率下，经典 + NLP 达到 10.01 EPE px，有力地比 PCA-流，更糟糕的是，需要 1.9 每帧对 s。PCA 层执行得更好，与地方 10 36 对单项传递，和对清洁传递 36 9。它执行特别好在无与伦比的地区，在那里它排列 5 36 两张通行证。这表明，我们学到的基础捕获的光流场结构不够好，大约只是一个框架的一双中可见的区域使“猜测”。叻。除了辛，我们测试我们叻基准 [16] 的方法。因为叻包含记录行驶中车辆的场景，

我们想到可能的运动是相对较低维子的空间。图 6 显示了一个示例。请注意我们是如何能够准确地判断运动的右侧对冲和如何的边界是更加锋利使用 PCA 层。在叻测试集，我们获取平均 EPE (Avg 全) 上所有像素的 6.2 PCA 流和 5.2 px px PCA 层。

在叻流纯粹是由汽车的低维空间的运动所造成的分割成层有助于更好地掌握运动边界。所有其他已发表的方法比 5 快每帧 s 表现差在平均 EPE，第二种最好是 TGV2CENSUS 与 6.6 px 在运行时的 4 s。与我们的类似精度没有基于 CPU 的方法是最快 10 s 每帧。

在出 Noc 度量 (百分比与错误 > 3 像素 px)、PCA 流量排名 40 63 (1567%) 和 PCALayers 排名 34 63 (12.02%)。这些结果 reflect 我们流场经久的近似性特点。

6.2. 计时

在当前 CPU (英特尔至强 i7), PCA 流算法平均需 190 ms 每帧在 MPI 辛数据集上。80 毫秒用于特征匹配。我们的算法的一个优点在于，当使用更长的序列，如那些从 MPI-辛，为每个帧的特征必须是只有一次，计算的消除坐标大约 20 毫秒运行时每帧。拟合的流量基础本身需要大约 90 毫秒。PCA 层是有力地速度较慢，平均需要 3.2 秒的镜架。我们的实现中使用 Python 和其 OpenCV 绑定。在 c++ 使用狢狢 [41] 实现了核心重组自交

系算法。图 7 地块辛和叻在 EPE 运行时平面 2 的最佳和最快发布方法。一般，所有方法比 PCA 流快需要 GPU，并实现更高的终点误差。另一方面，比 PCA 层更准确的所有方法都都有力地慢。

7. 结论和今后的工作

综上所述，本文进行了几个捐款。首先，我们证明计算全球光流场前瞻从大量的培训数据的基础的可行性。第二，我们表明此基础如何可以用不同的数据集和场景，表现出良好的泛化能力。第三，我们建议上有效估计近似的光流，一种算法使用稀疏特征匹配和学问的基础。第四，我们发展是更多比现有密集的分层方法规避稀疏层状光流场方法。为了这样做，我们结合几个 PCA 流经久使用图像证据来提高精度和产生急剧运动边界。第五，我们评价这两种算法对光流场估计的两个电流、挑战性数据集。我们的结果表明，疏到密方法可以对精度与竞争当前非稀疏方法同时实现先进的工件使用标准的 CPU。现有的基向量出现 sufficient 的任务和今后的工作应集中于“装配”相干流场前瞻从层 flows。我们当前的中期预测是很简单和更复杂的模型存在的场景分割。尤其是，包括高级别的场景来和分割成流场生成过程带来了希望。仍然，甚至在其当前的形式，结合的速度和准确性开辟了适用于涉及大型视频数据库的新问题的光流场的机会。

除了速度快，这里的紧凑性 PCA 流提供的光流场描述符也是互利，和我们目前正在探索 PCA 流用于大型视频来和索引。最后，我们的方法假设一个线性子空间。虽然这并不一定 reflect 流场的现实，与其他子空间提取方法的初步实验并没有改善精度。调查的子空间的光流场的真实结构仍然是未来的工作。

致谢。我们感谢 A. 盖革为从 [17] 的功能提供的代码。

参考：

- [1] <http://pcaflow.is.tue.mpg.de>. 2, 3, 4, 5, 7
- [2] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2(3):283 - 310, 1989. 2
- [3] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1 - 31, 2011. 1
- [4] L. Bao, Q. Yang, and H. Jin. Fast edge-preserving PatchMatch for large displacement optical flow. *Image Processing, IEEE Transactions on*, 23(12):4996 - 5006, Dec 2014. 2, 3

- [5] D. I. Barnea and H. Silverman. A class of algorithms for fast digital image registration. *Computers, IEEE Transactions on*, C-21(2):179 - 186, Feb 1972. 2
- [6] M. J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75 - 104, 1996. 5
- [7] M. J. Black and Y. Yacoob. Recognizing facial expressions in image sequences using local parameterized models of image motion. *International Journal of Computer Vision*, 25(1):23 - 48, 1997. 1
- [8] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11):1222 - 1239, Nov 2001. 6
- [9] T. Brox and J. Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(3):500 - 513, March 2011. 1, 2, 3, 4
- [10] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black. A naturalistic open source movie for optical flow evaluation. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, editors, *Computer Vision ECCV 2012*, volume 7577 of *Lecture Notes in Computer Science*, pages 611 - 625. Springer Berlin Heidelberg, 2012. 1, 2
- [11] Z. Chen, H. Jin, Z. Lin, S. Cohen, and Y. Wu. Large displacement optical flow from nearest neighbor fields. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2443 - 2450, June 2013. 2
- [12] M. Chessa, F. Solari, S. P. Sabatini, and G. M. Bisio. Motion interpretation using adjustable linear models. In *BMVC*, pages 1 - 10, 2008. 3
- [13] T. Darrell and A. Pentland. Robust estimation of a multi-layered motion representation. In *Visual Motion, 1991., Proceedings of the IEEE Workshop on*, pages 173 - 178, Oct 1991. 3
- [14] R. Fablet and M. Black. Automatic detection and tracking of human motion with a view-based representation. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Computer Vision - ECCV*

- 2002, volume 2350 of Lecture Notes in Computer Science, pages 476 – 491. Springer Berlin Heidelberg, 2002. 3
- [15] D. J. Fleet, M. J. Black, Y. Yacoob, and A. D. Jepson. Design and use of linear models for image motion analysis. *International Journal of Computer Vision*, 36(3):171 – 193, 2000. 1, 3, 4, 5
- [16] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research*, 32(11):1231 – 1237, Sept. 2013. 1, 2, 8
- [17] A. Geiger, J. Ziegler, and C. Stiller. StereoScan: Dense 3D reconstruction in real-time. In *Intelligent Vehicles Symposium (IV)*, 2011 IEEE, pages 963 – 968, June 2011. 4, 8
- [18] D. Gibson and M. Spann. Robust optical flow estimation based on a sparse motion trajectory set. *Image Processing, IEEE Transactions on*, 12(4):431 – 445, April 2003. 2
- [19] J. Gibson and O. Marques. Sparse regularization of TV-L1 optical flow. In A. Elmoataz, O. Lezoray, F. Nouboud, and D. Mammass, editors, *Image and Signal Processing*, volume 8509 of Lecture Notes in Computer Science, pages 460 – 467. Springer International Publishing, 2014. 3
- [20] T. Guthier, J. Eggert, and V. Willert. Unsupervised learning of motion patterns. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, volume 20, pages 323 – 328, Bruges, April 2012. 3
- [21] S. Hauberg, A. Feragen, and M. Black. Grassmann averages for scalable robust PCA. In *Computer Vision and Pattern Recognition (CVPR)*, 2014 IEEE Conference on, pages 3810 – 3817, June 2014. 1, 3
- [22] K. He, J. Sun, and X. Tang. Guided image filtering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(6):1397 – 1409, June 2013. 5
- [23] A. Jepson and M. Black. Mixture models for optical flow computation. In *Computer Vision and Pattern Recognition (CVPR)*, 1993 IEEE Conference on, pages 760 – 761, Jun 1993. 2, 3
- [24] R. Kennedy and C. Taylor. Optical flow with geometric occlusion estimation and fusion of multiple frames. In X.-C. Tai, E. Bae, T. Chan, and M. Lysaker, editors, *Energy Minimization Methods in Computer Vision and*

- Pattern Recognition, volume 8932 of Lecture Notes in Computer Science, pages 364–377. Springer International Publishing, 2015. 1, 2
- [25] P. Kr ¨ ahenb ¨ uhl and V. Koltun. Efficient nonlocal regularization for optical flow. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, editors, Computer Vision – ECCV 2012, volume 7572 of Lecture Notes in Computer Science, pages 356–369. Springer Berlin Heidelberg, 2012. 3
- [26] M. Lang, O. Wang, T. Aydin, A. Smolic, and M. Gross. Practical temporal consistency for imagebased graphics applications. ACM Trans. Graph., 31(4):34:1–34:8, July 2012. 1, 2
- [27] V. Lempitsky, S. Roth, and C. Rother. Fusionflow: Discrete–continuous optimization for optical flow estimation. In Computer Vision and Pattern Recognition (CVPR), 2008 IEEE Conference on, pages 1–8, June 2008. 6
- [28] M. Leordeanu, A. Zanfir, and C. Sminchisescu. Locally affine sparse–to–dense matching for motion and occlusion estimation. In Computer Vision (ICCV), 2013 IEEE International Conference on, pages 1721–1728, Dec 2013. 1, 2
- [29] C. Liu, J. Yuen, and A. Torralba. Sift flow: Dense correspondence across scenes and its applications. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 33(5):978–994, May 2011. 1, 2
- [30] J. Lu, H. Yang, D. Min, and M. Do. Patch Match Filter: Efficient edge–aware filtering meets randomized search for fast correspondence field estimation. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pages 1854–1861, June 2013. 2
- [31] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In IJCAI, volume 81, pages 674–679, 1981. 3, 5
- [32] O. Mac Aodha, A. Humayun, M. Pollefeys, and G. Brostow. Learning a confidence measure for optical flow. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35(5):1107–1120, May 2013. 6
- [33] M. Mohamed, H. Rashwan, B. Mertsching, M. Garcia, and D. Puig. Illumination–robust optical flow using a local directional pattern.

- Circuits and Systems for Video Technology, IEEE Transactions on, 24(9):1499 – 1508, Sept 2014. 2
- [34] M. Nicolescu and G. Medioni. Layered 4D representation and voting for grouping from motion. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 25(4):492 – 501, April 2003. 2
- [35] J. Rannacher. Realtime 3D motion estimation on graphics hardware. Undergraduate Thesis, Heidelberg University, 2009. 2
- [36] R. Roberts and F. Dellaert. Direct superpixel labeling for mobile robot navigation using learned general optical flow templates. In Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on, pages 1032 – 1037, Sept 2014. 3
- [37] R. Roberts, C. Potthast, and F. Dellaert. Learning general optical flow subspaces for egomotion estimation and detection of motion anomalies. In Computer Vision and Pattern Recognition (CVPR), 2009 IEEE Conference on, pages 57 – 64, June 2009. 3
- [38] D. Rosenbaum, D. Zoran, and Y. Weiss. Learning the local statistics of optical flow. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, editors, Advances in Neural Information Processing Systems 26, pages 2373 – 2381. MIT Press, 2013. 3
- [39] S. Roth and M. J. Black. On the spatial statistics of optical flow. International Journal of Computer Vision, 74(1):33 – 50, 2007. 3
- [40] C. Rother, V. Kolmogorov, and A. Blake. ” grabcut ” : Interactive foreground extraction using iterated graph cuts. ACM Trans. Graph., 23(3):309 – 314, Aug. 2004. 7
- [41] C. Sanderson. Armadillo: An open source C++ linear algebra library for fast prototyping and computationally intensive experiments. Technical report, NICTA, 2010. 8
- [42] T. Schoenemann and D. Cremers. A coding-cost framework for super-resolution motion layer decomposition. Image Processing, IEEE Transactions on, 21(3):1097 – 1110, March 2012. 3
- [43] D. Sun, C. Liu, and H. Pfister. Local layering for joint motion estimation and occlusion detection. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 1098 – 1105, June 2014. 1

- [44] D. Sun, S. Roth, and M. Black. A quantitative analysis of current practices in optical flow estimation and the principles behind them. *International Journal of Computer Vision*, 106(2):115 - 137, 2014. 2, 3
- [45] D. Sun, E. Sudderth, and M. Black. Layered segmentation and optical flow estimation over time. In *Computer Vision and Pattern Recognition (CVPR)*, 2012 IEEE Conference on, pages 1768 - 1775, June 2012. 2, 3
- [46] D. Sun, J. Wulff, E. Sudderth, H. Pfister, and M. Black. A fully-connected layered model of foreground and background flow. In *Computer Vision and Pattern Recognition (CVPR)*, 2013 IEEE Conference on, pages 2451 - 2458, June 2013. 1, 2, 3
- [47] N. Sundaram, T. Brox, and K. Keutzer. Dense point trajectories by gpu-accelerated large displacement optical flow. Technical Report UCB/EECS-2010-104, EECS Department, University of California, Berkeley, Jul 2010. 3
- [48] M. Tao, J. Bai, P. Kohli, and S. Paris. Simpleflow: A non-iterative, sublinear optical flow algorithm. *Computer Graphics Forum*, 31(2pt1):345 - 353, 2012. 3
- [49] J. Wang and E. Adelson. Representing moving images with layers. *Image Processing, IEEE Transactions on*, 3(5):625 - 638, Sep 1994. 3
- [50] P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid. DeepFlow: Large displacement optical flow with deep matching. In *Computer Vision (ICCV)*, 2013 IEEE International Conference on, pages 1385 - 1392, Dec 2013. 1, 2, 3, 4
- [51] M. Werlberger, T. Pock, and H. Bischof. Motion estimation with non-local total variation regularization. In *Computer Vision and Pattern Recognition (CVPR)*, 2010 IEEE Conference on, pages 2464 - 2471, June 2010. 3
- [52] M. Werlberger, W. Trobin, T. Pock, A. Wedel, D. Cremers, and H. Bischof. Anisotropic Huber-L1 optical flow. In *BMVC*, London, UK, September 2009. 1, 2, 3
- [53] L. Xu, J. Jia, and Y. Matsushita. Motion detail preserving optical flow estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(9):1744 - 1757, Sept 2012. 1, 2, 4

[54] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime TV-L1 optical flow. In F. Hamprecht, C. Schnrr, and B. Jhne, editors, Pattern Recognition, volume 4713 of Lecture Notes in Computer Science, pages 214 - 223. Springer Berlin Heidelberg, 2007. 2, 8

[55] K. Zuiderveld. Contrast limited adaptive histogram equalization. In P. S. Heckbert, editor, Graphics Gems IV, pages 474 - 485. Academic Press Professional, Inc., San Diego, CA, USA, 1994. 4