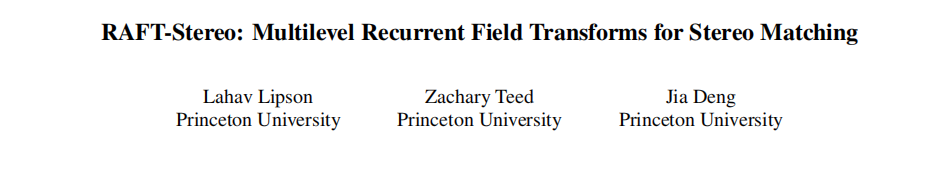
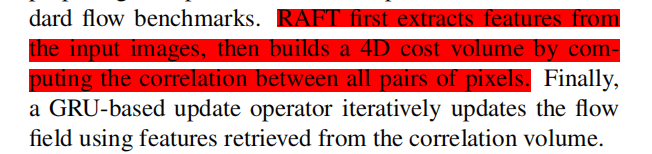
本周阅读了文章：



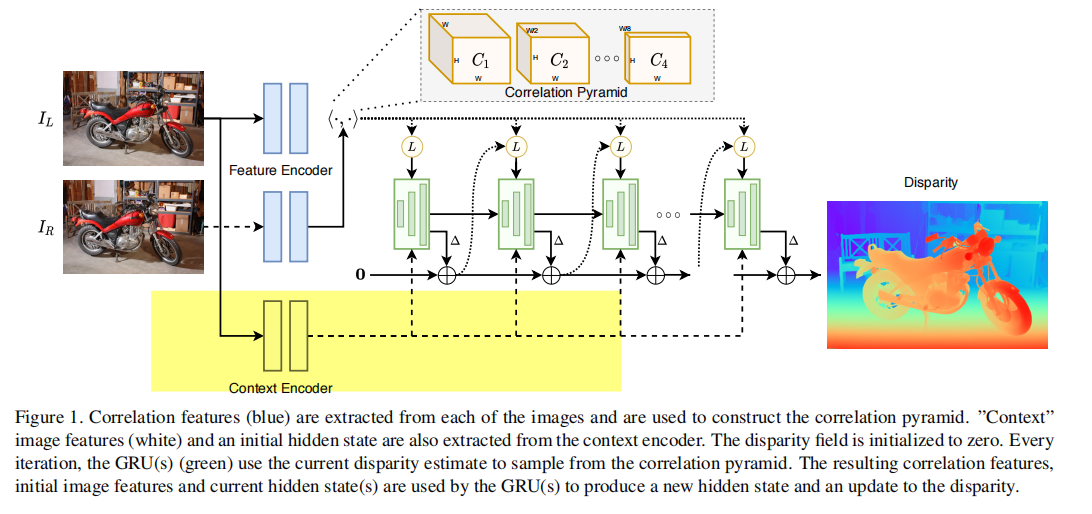
总体的阅读和对文章思路的掌握难度比上周要简单，一方面是这次认真的阅读了文章综述，另一方面是这篇文章介绍背景（”optical flow”和”rectified stereo”这两个任务在上周的文章也有所提到，感觉应该先阅读这篇文章再阅读上周的文章这个顺序才是正确的）

这篇文章基于原有的RAFT方法进行了结构的修改和优化，从而获得了文章提出的RAFT-Stereo方法，文章在综述部分介绍了stereo领域的发展（文章中指出了3D convolutional neural networks的不足：通用性差（一任务一方法、较大的计算成本）），同时指出optical flow方面的RAFT的基本思路：



提出的RAFT-stereo方法的优势在于，与以往的方法不同的是，该方法基于2D卷积（通过）和轻量级矩阵乘法构造cost volume。（能够有效节约计算资源和时间）

模型架构介绍



文章将模型分成了三个部分进行介绍（主要介绍文章中强调“值得指出的部分”）：

**Feature Extraction**

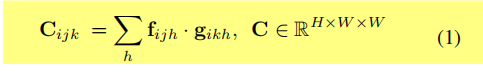
文章使用两个特征提取器：

The feature extractor：文章中的相关参数设置提及的较为详细。

The context extractor：在架构中用于初始化更新操作符的隐藏状态，并在更新操作符的每次迭代中被嵌入到GRU中。

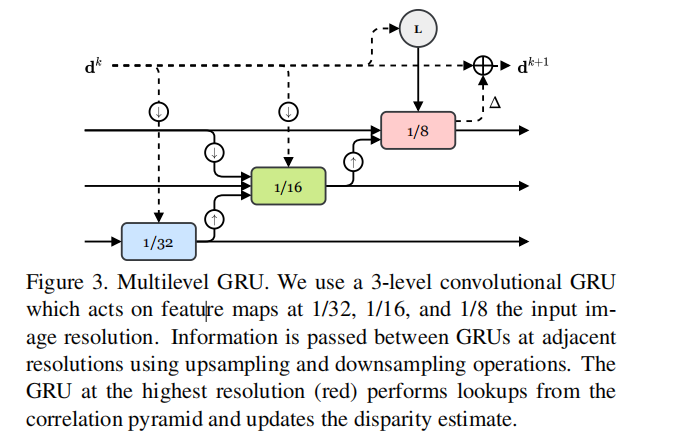
**Correlation Pyramid**

围绕当前视差估计值构造一个具有整数偏移量的一维网格+整数偏移量上的线性插值（采用三维volume表示，利于矩阵乘法实现）



**GRU based Update Operator**

和原有的RAFT方法有所不同的一点在于原有的方法使用在固定的、high resolution下执行更新（对大段拥有较少局部信息的情况来说存在问题），文章引入了GRU来代替上述方法



GRU模块说明如上（在后面的实现中有提到1/8GRUhidden state的浮点数运算量是计算1/16的4倍，RAFT-Stereo采用多次更新1/16和1/32分辨率的隐藏状态，并每次更新到1/8分辨率的隐藏状态用于减少运行时间）同时每个hidden state都使用专用GRU进行更新，该GRU除了使用解析的特定上下文特性外，还使用相邻的hidden state作为上下文来进行上下文。

损失函数采用质数型增长权重设计：



优化器采用adamW优化器实现，具体实践细节参数选择（P5的traisning schedule给出了）