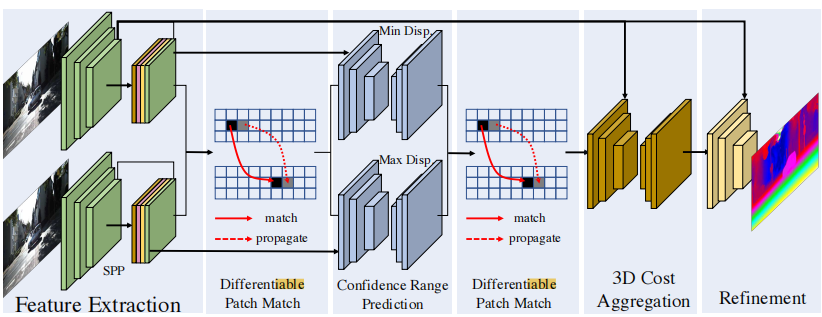
论文目标：基于可微patch match模块的立体匹配算法优化

做了什么：提出了一个可微的PatchMatch模块，可以在没有完整成本体积评估的情况下丢弃大部分视差，并利用稀疏表示来学习为每个像素剪枝的范围。

效果:我们能够有效地计算高似然假设的成本量，并实现在内存和计算方面的节省。

实现思路：

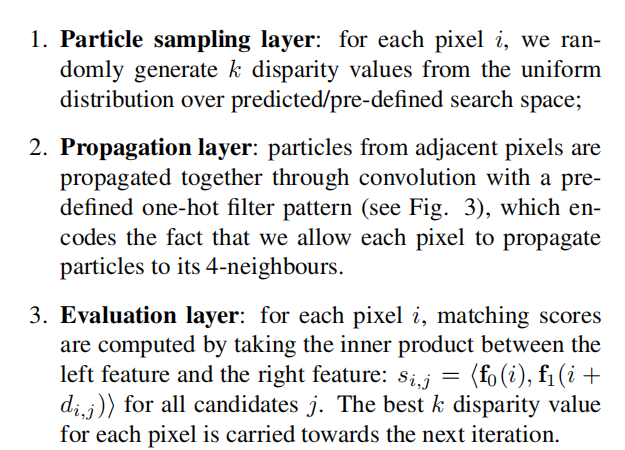


初始阶段

首先，模型对左右图像进行特征提取，在特征图上为每个像素预测一个初始的视差估计。

具体来说：作者从输入图像中提取可靠的像素级特征表示。为了实现这一点，作者采用了一个具有空间金字塔池化模块的二维卷积神经网络作为基础。网络包含4个残差块，并且在最后一个块中采用×2空洞卷积来扩大感受野。接下来，作者应用空间金字塔池化来构建一个4级金字塔特征。通过多尺度信息，模型能够捕捉到大的上下文并保持高的空间分辨率。最终的特征图大小是原始图像尺寸的1/4。作者共享左右特征网络的参数。

Patch match过程：



·初始化：首先，模型使用随机采样的方法生成k个候选视差值，这些候选视差值以均匀分布在预测/预定义的搜索空间上。

·传播（Propagation）：接下来，模型通过将相邻像素的粒子进行一起传播来执行传播操作。具体来说，模型使用预定义的一种卷积核将粒子从相邻像素进行传播，以此来将信息从邻居像素传递到当前像素。

·评估（Evaluation）：在评估阶段，模型通过计算左侧特征和右侧特征的内积来计算每个像素的匹配分数。对于每个像素i，匹配分数由公式：

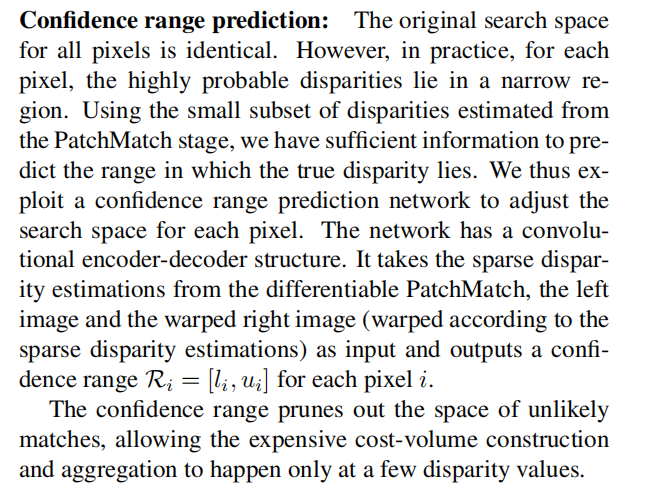


计算得到，其中di,j是第j个候选视差值。模型选择前k个匹配得分最高的候选视差值作为每个像素的最佳视差值，并将其传递到下一次迭代。

通过将经典的PatchMatch算法转化为可微分形式的思路是为了解决传统立体匹配中大规模搜索空间带来的内存和计算负担问题，该模块能够为每个像素预测一个置信区间，并构建一个稀疏的成本体积，从而大大减少了计算量和内存消耗。将其设计为递归神经网络的一部分。从而算法的迭代过程被划分为一系列的网络层，每个网络层都等价于PatchMatch算法的一次迭代。这样，整个算法可以通过网络的反向传播进行训练，并且可以端到端地进行优化。

可微PatchMatch模块

为了实现端到端的学习，作者将通用的PatchMatch算法展开为一个循环神经网络。每个循环步骤对应着一次算法迭代。在每个循环步骤中，模型通过粒子的采样、传播和评估来逐步更新视差估计。



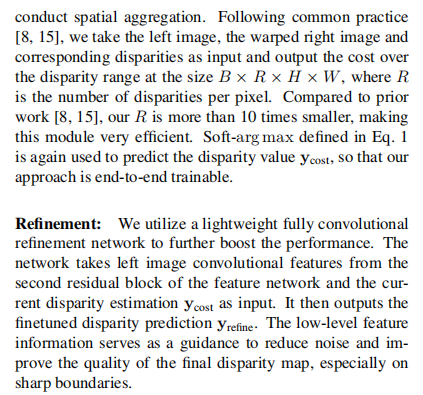
置信区间预测网络

基于修剪模块中的稀疏视差估计结果，作者设计了一个置信区间预测网络，用于预测每个像素的置信区间。该网络接受稀疏视差估计、左图像和经过视差估计进行了扭曲的右图像作为输入，并输出每个像素的置信区间。

成本聚合和优化

根据置信区间的预测结果，模型构建了一个三维成本体积，用于立体匹配。在此基础上，模型进行成本聚合和优化，以进一步提高视差估计的准确性。

具体来说：



成本聚合

根据置信区间的预测结果，作者建立了一个稀疏的成本体积估计器。

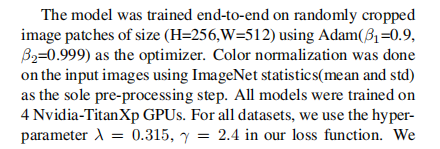
稀疏的成本体积只计算了置信度较高的视差值，将计算和存储资源的开销大大降低。作者利用置信度范围剪枝了视差空间的不可能匹配的部分，从而只对几个视差值进行了计算。这种方法有效地减少了计算量和内存消耗，同时保证了性能。

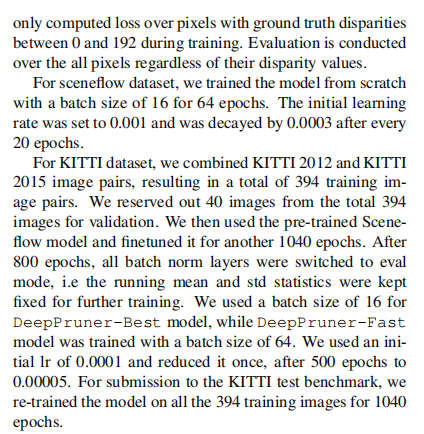
优化

为了进一步优化视差估计结果，作者将左侧图像、经过稀疏视差估计的右侧图像以及稀疏视差估计结果作为输入，并生成更精确的视差估计结果。通过利用图像的上下文信息，这个模块可以纠正稀疏视差估计的错误，并提高整体的视差估计质量。

在整个成本聚合和优化的过程中，作者使用了可微patch match模块，使得整个网络可以进行端到端的训练。这意味着可以通过最小化估计视差和真实视差之间的差异来优化网络的参数，从而得到更精确的视差估计结果。

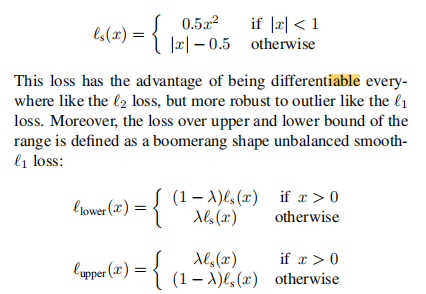
实现细节





该部分主要介绍参数对不同图像集合的参数设置（如学习率、权重设计）

同时提出了损失函数设计：



效果评估

作者使用了KITTI 2012和KITTI 2015数据集进行评估，这些数据集是常用的用于视觉算法评估的数据集。 评估过程中，作者对比了DeepPruner和其他几种先进的实时深度学习立体匹配模型，包括PSMNet、GANet和GC-Net。作者还使用了不同的评测指标，包括Bad 3-pixel、D1-all和D2全部，这些指标用于度量立体匹配的精度和准确性。 实验结果表明，DeepPruner在运行时间和准确性上均取得了显著的改进。与其他模型相比，DeepPruner在运行时间上实现了真实时间推理，每个图像的运行时间仅为62ms。同时，在精度方面，DeepPruner在各项指标上都取得了竞争性的结果，比其他模型更接近最先进的方法。

评估结果表明，DeepPruner是一种高效且准确的实时深度学习立体匹配模型，适用于各种视觉应用，如机器人导航、自动驾驶和增强现实等。