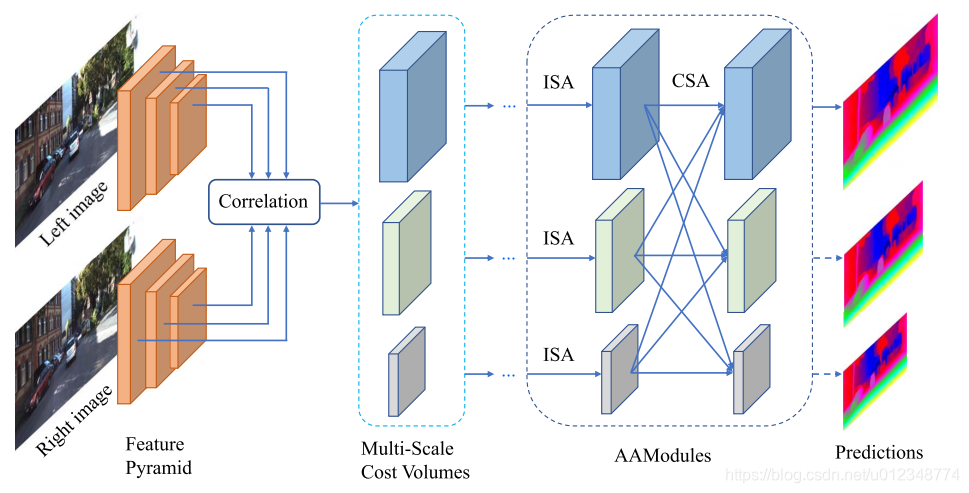
AANet:



input: 设置最大视差为192

left, right 1\*3\*312\*420

backbone out:

f1: 1\*128\*104\*140

f2: 1\*256\*52\*70

f3: 1\*512\*26\*35

主干网络得到3个尺寸,分别是原图的1/3, 1/6, 1/12

FPN out:

fp1: 1\*128\*104\*140

fp2: 1\*128\*52\*70

fp3: 1\*128\*26\*35

分辨率最小的放大后与上一层相加

fp3 = f3, fp2 = fp3\*2 + f2, fp3 = fp2 \* 2 + f1

代价计算out:

fp1: 1\*64\*104\*140

fp2: 1\*32\*52\*70

fp3: 1\*16\*26\*35

代价计算为左右特征图的相关运算结果, 通道维度为对应的视差, 64, 32, 12为原图最大视差缩放1/3, 1/6, 1/12到对应尺寸后的最大视差

代价聚合:

fp1: 1\*64\*104\*140

fp2: 1\*32\*52\*70

fp3: 1\*16\*26\*35

首先先对每个尺寸自身做卷积运算做自聚合, 然后再与其他两个尺寸做交叉聚合,即卷积再缩放再相加

视差计算:

对每个尺寸在通道/视差方向做softmax,然后再加权求和,权重就是softmax的结果,再乘上对应视差值

1\*26\*35

1\*52\*70

1\*104\*140

视差精炼:

用最大分辨率的视差(1\*104\*140)逐渐回归回原图尺寸(1\*312\*420)