

# 2019-08-07-郭叙森 周报

对这两周实验总结如下：

## 物体检测指标

下表是目前做的几个实验对比，IOU=0.5 表示当预测框与真实框交并比不小于0.5时记作检测正确，IOU=0.7 则交并比阈值是0.7. 每一项分别为  $BEV_{AP}/3D_{AP}$ ，单位为%，分别表示在俯视图上的评估（只考虑俯视图上的重叠情况）以及3D上的评估（三维框的重叠情况）。Easy/Moderate/Hard 是 KITTI 数据集对目标检测困难程度的分类，Easy 一般是能够很好的观测到，没有遮挡或者遮挡很少，而 Moderate 与 Hard 则表示遮挡逐渐增大，目标不好辨识。Stride 表示插值的步长，1表示两关键帧编号相差1，即是相邻关键帧，这时候不需要插值。-1- 表示插值过程中不对轨迹的首尾两端进行延长，+1+ 表示插值过程中对轨迹首尾进行延长。本组实验延长的帧数都是5帧。以下是实验数据：

-	IOU = 0.5			IOU = 0.7		
Stride	Easy	Moderate	Hard	Easy	Moderate	Hard
1	90.59/99.98	88.94/90.90	88.76/90.89	88.33/90.90	75.25/90.86	68.51/90.85
-1-	84.83/99.98	83.15/90.89	83.68/90.06	78.77/90.81	65.15/90.77	59.35/90.76
+1+	82.84/99.91	81.36/90.84	81.98/90.84	76.28/90.76	63.49/90.68	57.99/90.68
-2-	79.53/90.82	77.38/81.75	77.46/81.74	75.85/81.73	62.95/81.65	63.18/81.65
+2+	82.23/90.57	73.86/81.56	74.20/81.58	71.23/90.40	65.49/81.46	60.13/81.49

参数设置（备忘）： $\sigma_l = 0.1, \sigma_h = 0.3, \sigma_{iou} = 0.1, t_{min} = 3, extend_{len} = 5$

从以上结果可以看出，不做任何处理的检测结果“1”对比插值算法方案但不做任何轨迹延长“-1-”效果要好大概10%（IOU=0.7），这里的差距需要格外注意，我们的算法在这里应该有很大的提升空间。原因初步分析如下：由于stride是1，没有插值过程的引入，因此问题主要出现在匹配上。由于新算法在相邻帧物体相匹配且持续一定时间才认为是有效物体，因此算法对目标跳动比较大的情况下检测不佳，会认为是噪声。还应有其他原因，具体问题还在探究。

“-1-”与“+1+”对比可发现添加轨迹首尾会使性能下降，原因是KITTI原始标签并没有对所有物体打上标签。对于那些在图像平面上不存在的（但是点云中有），或者很小的目标，并没有真值标签。但是我们的算法根据运动估计轨迹时会根据物体的运动生成标签，很多生成的三维框事实上是有用的只是KITTI没有打上标签，才使得准确率下降。

“+1+”与“+2+”对比可以发现，当IOU=0.7时，“+2+”在 Moderate 与 Hard 类别上要比“+1+”好，这得益于时序信息的加入。因为 Moderate 与 Hard 类别 大多是遮挡的物体，时序信息的引入有利于这些物体的检测。而在Easy 类别上的性能下降可能是由于插值过程中目标朝向的不准确，导致插值不准确。

## 多目标追踪对比

以下是多目标跟踪实验数据，其中 Method 一栏和目标检测的 Stride 一样。先介绍追踪常用的评估指标：

**MOTA**：Multiple Object Tracking Accuracy,  $MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t}$ ，其中  $m = \frac{\sum_t m_t}{\sum_t g_t}$  计算的是 tracking过程中总体的漏检率， $fp = \frac{\sum_t fp_t}{\sum_t g_t}$  计算的是tracking过程总体的误检，也就是虚警率， $mme = \frac{\sum_t mme_t}{\sum_t g_t}$  表示tracking中所有目标发生跳变的几率。

**MOTP**：Multiple Object Tracking Precision,  $MOTP = \frac{\sum_{i,t} d_t^i}{\sum_t c_t}$ ，也就是假设与对应目标的平均偏差。

**MT**: Mostly Tracked. 指跟踪部分大于80%的跟踪轨迹数，数值越大越好。

**ML:** Mostly Lost. 指丢失部分大于80%的跟踪轨迹数，数值越小越好。

**IDS:** ID-Switches, 跟踪轨迹中目标 ID 瞬间转换的次数，通常能反应跟踪的稳定性，数值越小越好。

**FRAG:** Fragmentation, 指跟踪轨迹从“跟踪”到“不跟踪”状态的变化数。

**Recall：**召回率是指被分类器正确判定的正例占总的正例的比重。

**Precision：**精确度是指被分类器判定的正例中真正的正例样本的比重。

Method	MOTA(%)	MOTP(%)	MT(%)	ML(%)	IDS(%)	FRAG	Recall(%)	Precision(%)
1	76.73	84.47	65.55	6.30	15	121	80.57	97.73
-1-	78.11	83.82	68.49	5.89	22	85	84.50	95.18
+1+	77.01	83.60	72.27	5.89	30	91	85.56	93.37
-2-	73.11	84.58	60.92	13.02	37	124	77.98	96.73
+2+	71.00	83.96	67.65	12.61	64	128	80.69	92.33

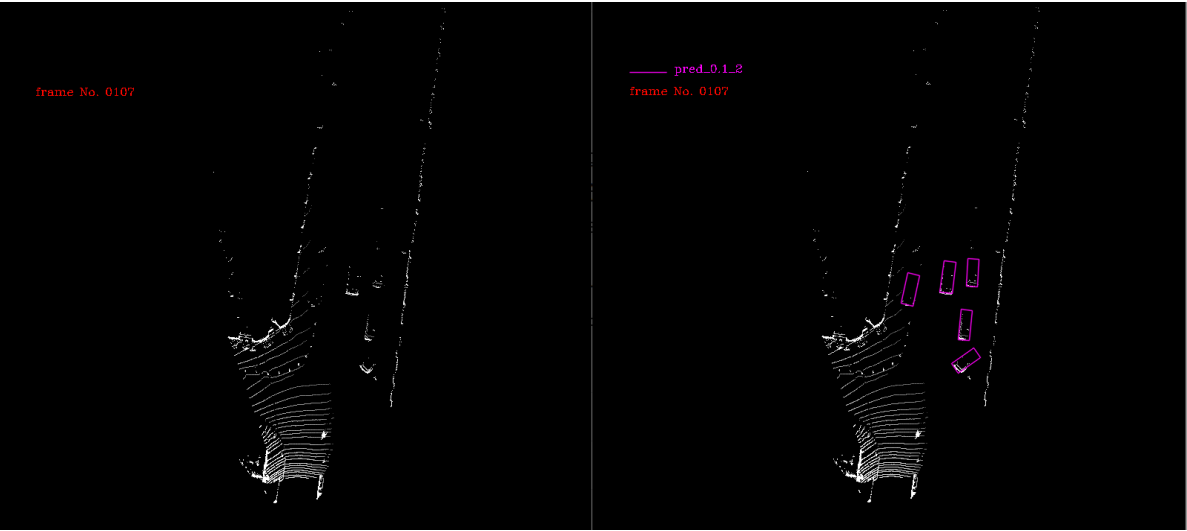
多目标追踪结果如上表所示。

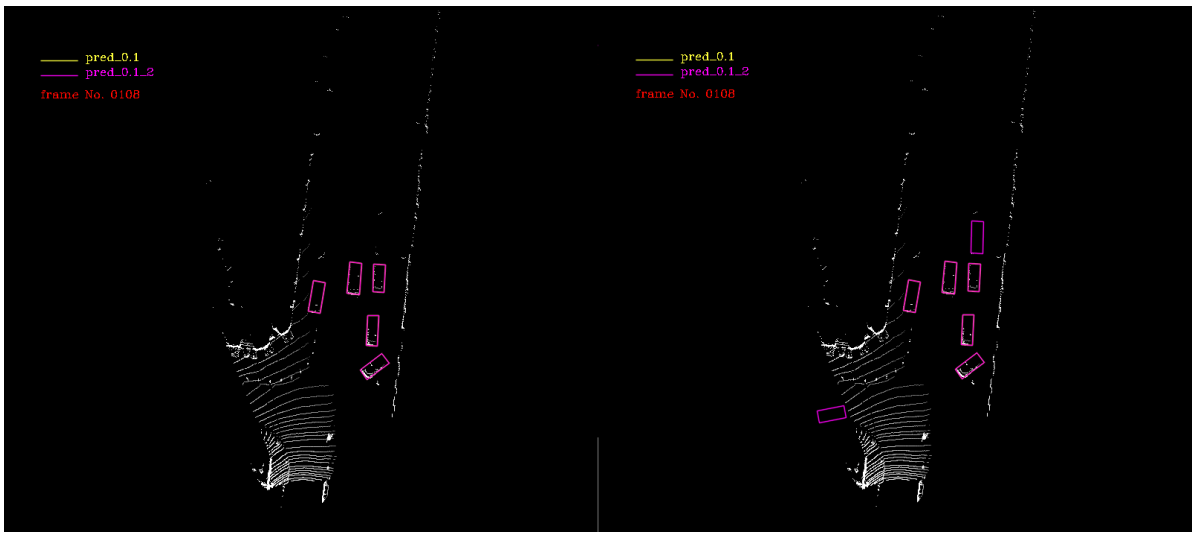
由“1”和“-1-”对比可知，我们的算法能够提升多目标跟踪的性能，这应该得益于时序信息的加入。注意到“-1-”召回率 Recall 升高但 精确率 Precision 降低，这就印证了上面说的我们的方法能够检测出更多的真值框，但由于KITTI数据集标签的不全面，导致准确率下降。

由“-1-”与“+1+”对比可知，轨迹首尾的补全提升了召回率，但是精确率下降。这更加正是了上一条的猜想。另外值得注意的是，加入首尾后 MT 显著增大，这是因为轨迹更加完善了。

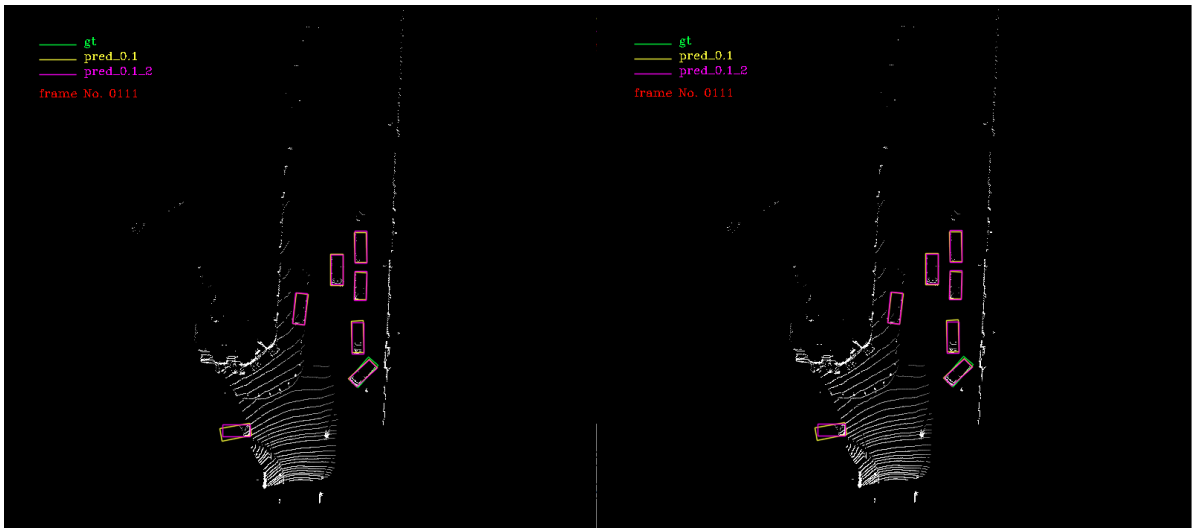
由“+1+”与“+2+”对比可知，插值算法并不能很好的工作，对目标跟踪的影响比较大，还需继续完善。

图示说明





上图是连续两帧，左边是“-2-”，右边是“+2+”，也即是右边有进行轨迹首尾补全。可以发现第一帧左边没有目标，真实标签也没有，但实际上我们肉眼可以发现其实是有目标的，从后几帧的数据（下图）也能看出是有的。这说明的确存在上面分析的情况，即真实标签的不完善导致我们的结果准确率下降。



## 下周计划

进一步分析结果不佳的原因，试试卡尔曼滤波看看能否解决目标朝向不准确的问题。