

一类自适应的数据提升方法

摘要. 本文提出了一套数据评估和提升的整套算法, 包括自评估, 联合评估和提升. 文章主要分 5 个部分, 分别是图片自提升, 点云自提升, 图片点云数据联合校准, 和理论证明, 实验验证

1. 引言

在自动驾驶技术中, 我们通常会遇到这样的问题, 如果我们采用多传感器融合的方式, 有时候会发现, 单个数据集进行训练反而要比多传感器融合的效果更差. 比如: 这不是我们希望得到的结果. 因此, 本文提出了一整套流程, 来解决这个问题.

2. 图像数据的评估和提升

2.1. 焦盐噪声.

2.1.1. 评估算法.

2.1.2. 提升算法.

2.2. 曝光异常.

2.2.1. 评估算法.

2.2.2. 提升算法.

2.3. 运动模糊.

2.3.1. 评估算法.

2.3.2. 提升算法.

3. 点云数据的评估和提升

3.1. 点云稀疏.

4. 融合评估与联合校准

4.1. 评估算法, 可能用到的数学方法

4.2. 校准算法.

5. 理论证明

这一节我们将用数学的方法严格证明我们所采用的方法可以达到一定的效果。

5.1. 假设为正太分布的合理性. 一方面, 根据中心极限定理, 正太分布是任意分布的极限分布。所以, 对我们不知道任何信息的分布, 一般都是假设为正太分布。另一方面, 任意分布, 都可以唯一地对应到一个特征函数。而均值和方差和这个特征函数的一阶导数和二阶导数有关。所以, 我们对特征函数泰勒展开 (如果存在), 知道一阶项和二阶项, 就可以假设为正太分布。

5.2. 数据提升后输入数据有更小的方差.

5.2.1. 去除奇异值对方差的影响.

5.2.2. 卡尔曼滤波器对方差的影响.

5.3. 数据提升后具有更好的鲁棒性. 设我们的输入为 $x_1 = (Image, Points)$, $g(x_1) = x_2$, 且 $D(x_1) > D(x_2)$. 然后一般的, 如果 $f(x)$ 有任意价的, 我们有 $f(x) = x_0 + f'(x_0)x + \frac{1}{2!}f''(x_0)x^2 + \frac{1}{3!}f^3(x_0)x^3 + \dots$ ¹

$$\begin{aligned}
 (1) \quad D(f(x)) &= D(x_0 + f'(x_0)x + f''(x_0)x^2 + f^3(x_0)x^3 + \dots), \\
 &= f'(x_0)D(x) + f''(x_0)D(x^2) + f^3(x_0)D(x^3) + \dots, \\
 &= (f'(x_0) + a_2f''(x_0) + a_3f^3(x_0) + \dots)\sigma^2, \\
 &= K\sigma^2
 \end{aligned}$$

当 k 为奇数时,

$$(2) \quad D(x^k) = (2k-1)!!\sigma^2$$

当 k 为偶数时,

$$(3) \quad D(x^k) = ((2k-1)!! - (k-1)!!)\sigma^2$$

所以 $D(f(x_1)) < D(f(x_2))$.

¹师兄, 这儿我想先证明在单变量时候的结论, 然后推广到多变量, 然后推广到大维统计

5.4. **自适应参数的选择.** 一般的,我们考虑更两个随机变量 $x_1 \sim N(a_1, \sigma_1^2), x_2 \sim N(a_2, \sigma_2^2), x_1, x_2$ 相互独立.² 则 $(x_1, x_2) \sim N(a_1, \sigma_1^2; a_2, \sigma_2^2; 0)$. 令 $y = \lambda_0 + \lambda_1 x_1 + \lambda_2 x_2$. 则

$$\begin{aligned} \phi_y(t) &= e^{j\lambda_0 t} \phi(\lambda_1 t, \lambda_2 t) \\ &= e^{j\lambda_0 t} e^{j(\lambda_1 a_1 t + \lambda_2 a_2 t)} \\ &\quad - \frac{1}{2}(\sigma_1^2 \lambda_1^2 t^2 + \sigma_2^2 \lambda_2^2 t^2) \end{aligned} \quad (4)$$

$y \sim N(\lambda_0 + \lambda_1 a_1 + \lambda_2 a_2, \sigma_1^2 \lambda_1^2 + \sigma_2^2 \lambda_2^2)$. 所以根据上一节的讨论, 我们要输出有更好的稳健性, 只需要线性组合 $y = a_1 x_1 + a_2 x_2$ 有更小的方差即可, 也就是极小化 $\lambda_1^2 \sigma_1^2 + \lambda_2^2 \sigma_2^2$. 根据三角不等式 $\lambda_1^2 \sigma_1^2 + \lambda_2^2 \sigma_2^2 \geq 2\lambda_1 \lambda_2 \sigma_1 \sigma_2$, 当 $\lambda_1 \sigma_1 = \lambda_2 \sigma_2$ 时取等号。

5.5. **自适应参数的选择.** 一般的,我们考虑更两个随机变量 $x_1 \sim N(a_1, \Sigma_1^2), x_2 \sim N(a_2, \Sigma_2^2), x_1, x_2$ 相互独立. 则 $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}\right)$. 令 $y = \lambda_1 x_1 + \lambda_2 x_2$. 则

$$\lambda_1 x_1 + \lambda_2 x_2 = (\lambda_1, \lambda_2) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (5)$$

故

$$E(\lambda_1 x_1 + \lambda_2 x_2) = (\lambda_1, \lambda_2) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (6)$$

$$D(\lambda_1 x_1 + \lambda_2 x_2) = (\lambda_1, \lambda_2) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (7)$$

$y \sim N\left((\lambda_1, \lambda_2) \begin{pmatrix} f a_1 \\ a_2 \end{pmatrix}, (\lambda_1, \lambda_2) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}\right)$. 所以根据上一节的讨论, 我们要输出有更好的稳健性, 只需要线性组合 $y = a_1 x_1 + a_2 x_2$ 有更小的协方差矩阵即可, 也就是极小化 $(\lambda_1, \lambda_2) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$.

$$\text{而 } (\lambda_1, \lambda_2) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \lambda_1^T \Sigma_1 \lambda_1 + \lambda_2^T \Sigma_2 \lambda_2.$$

如果我们取矩阵的某个范数来度量协方差矩阵的大小, 那么我们可以相应的得到我们的权重向量 λ_1 和 λ_2 . 不妨我们取矩阵的最大特征值作为范数. 那么我们可以证明当 $\lambda_1 \sigma_1 = \lambda_2 \sigma_2$ 时取到最小值. 其中 σ_1^2 是 Σ_1 的最大特征值, σ_2 是 Σ_2 的最大特征值.

²师兄, 这儿也是先证明单边量, 然后证明多变量, 然后证明大维度, 如果后面做出来我会把这部分删掉。

6. 仿真实验

6.1. KITTI 数据集. KITTI 数据集由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创办，是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集。该数据集用于评测立体图像 (stereo)，光流 (optical flow)，视觉测距 (visual odometry)，3D 物体检测 (object detection) 和 3D 跟踪 (tracking) 等计算机视觉技术在车载环境下的性能。KITTI 包含市区、乡村和高速公路等场景采集的真实图像数据，每张图像中最多达 15 辆车和 30 个行人，还有各种程度的遮挡与截断。整个数据集由 389 对立体图像和光流图，39.2 km 视觉测距序列以及超过 200k 3D 标注物体的图像组成，以 10Hz 的频率采样及同步。

6.2. 测试结果.

6.2.1. 准确性.

6.2.2. 鲁棒性.

6.2.3. 时间.

7. 结论

³在实际操作当中，我们得到一张照片和点云没有办法判断方差的大小，所以，可以利用拍到的 7 张照片或点云作为一个小样本，用前面的 3 个，后面的 3 组作为一个样本组估计出方差，这样做的合理性在于如果频率很高的相机和点云相机，可以近似的看成是对同一物体的测量。