

单目视觉是Mobileye (ME) 的看家法宝，其实当年它也考虑过双目，最终选择放弃。

单目的测距和3-D估计靠什么？是检测目标的Bounding Box (BB)，如果无法检测的障碍物，该系统就无法估计其距离和3-D姿态/朝向。没有深度学习的时候，ME主要是基于BB，摄像头标定得到的姿态和高度以及路面平直的假设估算距离。

有了深度学习，可以根据3-D的ground truth来训练NN模型，得到3D大小和姿态估计，距离是基于平行线原理 (single view metrology) 得到的。不久前百度Apollo公布的单目L3解决方案讲的比较清楚了，参考论文是"3D Bounding Box Estimation by Deep Learning and Geometry".

双目当然可以算视差和深度了，即使没有检测出障碍物（因为有附加的深度信息，检测器会比单目好），也会报警。问题是，双目视觉系统估计视差没那么容易，立体匹配是计算机视觉典型的难题，基线宽得到远目标测距准，而基线短得到近目标测距结果好，这里是存在折衷的。

目前市场上ADAS存在的双目视觉系统就是Subaru EyeSight，据说性能还行。



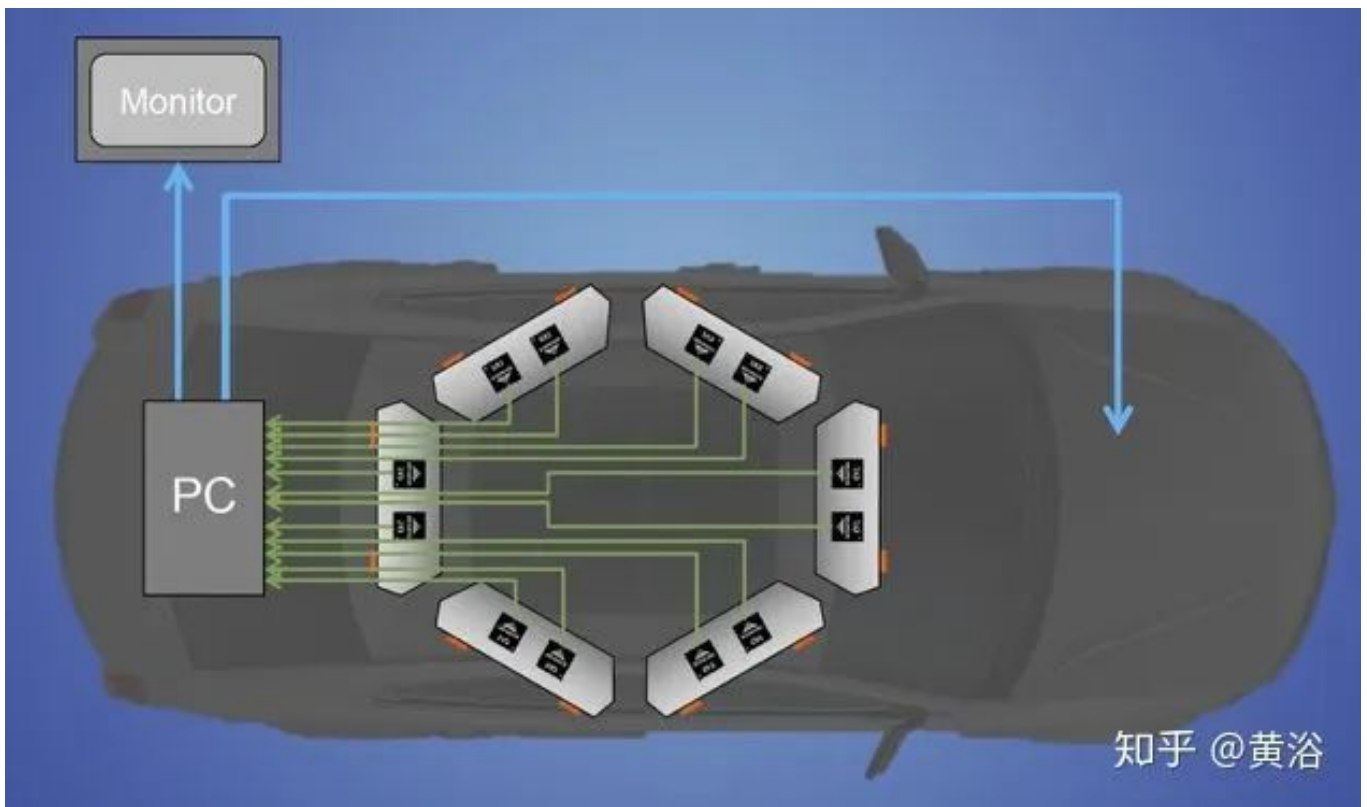
百度推出的阿波龙L4摆渡车量产100台，就安装了双目系统。还有欧盟自主泊车项目V-Charge也采用了前向双目视觉系统，另外自动驾驶研发系统Berta Benz也是，而且和雷达系统后融合，其中双目匹配的障碍物检测算法Stixel很出名。以前Bosch和Conti这些Tier-1公司也研制过双目视觉解决方案，但没有在市场上产生影响力，据说被砍掉了。

谈到双目系统的难点，除了立体匹配，还有标定。标定后的系统会出现“漂移”的，所以在线标定是必须具有的。单目也是一样，因为轮胎变形和车体颠簸都会影响摄像头外参数变化，必须在线做标定修正一些参数，比如仰角（pitch angle）和偏角（yaw angle）。

双目在线标定就更复杂些，因为双目匹配尽量简化成1-D搜索，所以需要通过stereo rectification将两个镜头光轴方向平行并和基线垂直。所以针对获得的gain相比，增加的复杂度和成本，如果不划算商家就会放弃。

最近重提双目视觉，是因为硅谷芯片公司安霸（Ambarella）在2014年收购意大利帕尔马大学的Vis Lab，研制了双目的ADAS和自动驾驶芯片，去年CES之后就开始进军车企和Tier-1。而且，安霸目前正在继续研究提升该系统的性能。

下图就是它在车顶安装6对立体视觉系统的示意图，其中它们的基线宽度可以不一样的，相应地有效检测距离也就不同。笔者曾坐过它的自动驾驶车，远处可以看到200米，近处20-30米。它确实可以做在线标定，随时调整一些双目视觉的参数。



先说立体匹配。

匹配方法分两种，全局法和局部法，实用的基本是局部法，因为全局法太慢。深度学习可以取代全局法，但目前还不是很成熟。

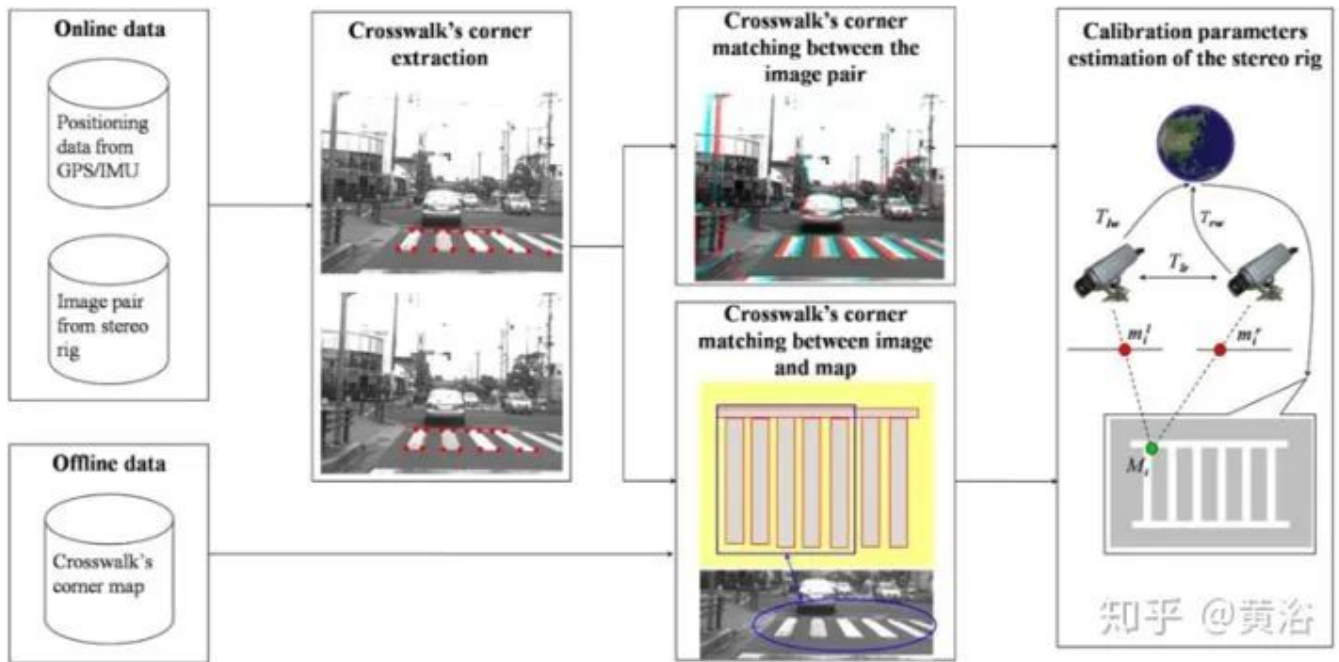
最著名的局部法就是SGM（semi-global matching），很多产品在用的方法都是基于此的改进，不少视觉芯片都采用这种算法。

Census Transform是将8/24比特的像素变成一个2进制序列，另外一个2值特征叫LBP（local binary pattern）和它相似。立体匹配算法就是基于这个变换将匹配变成一个Hamming距离的最小化搜索。Intel的RealSense当年就是收购了一个成立于1994年基于该技术的双目视觉创业公司，还收购另外几个小公司把他们合在一起做出来的。

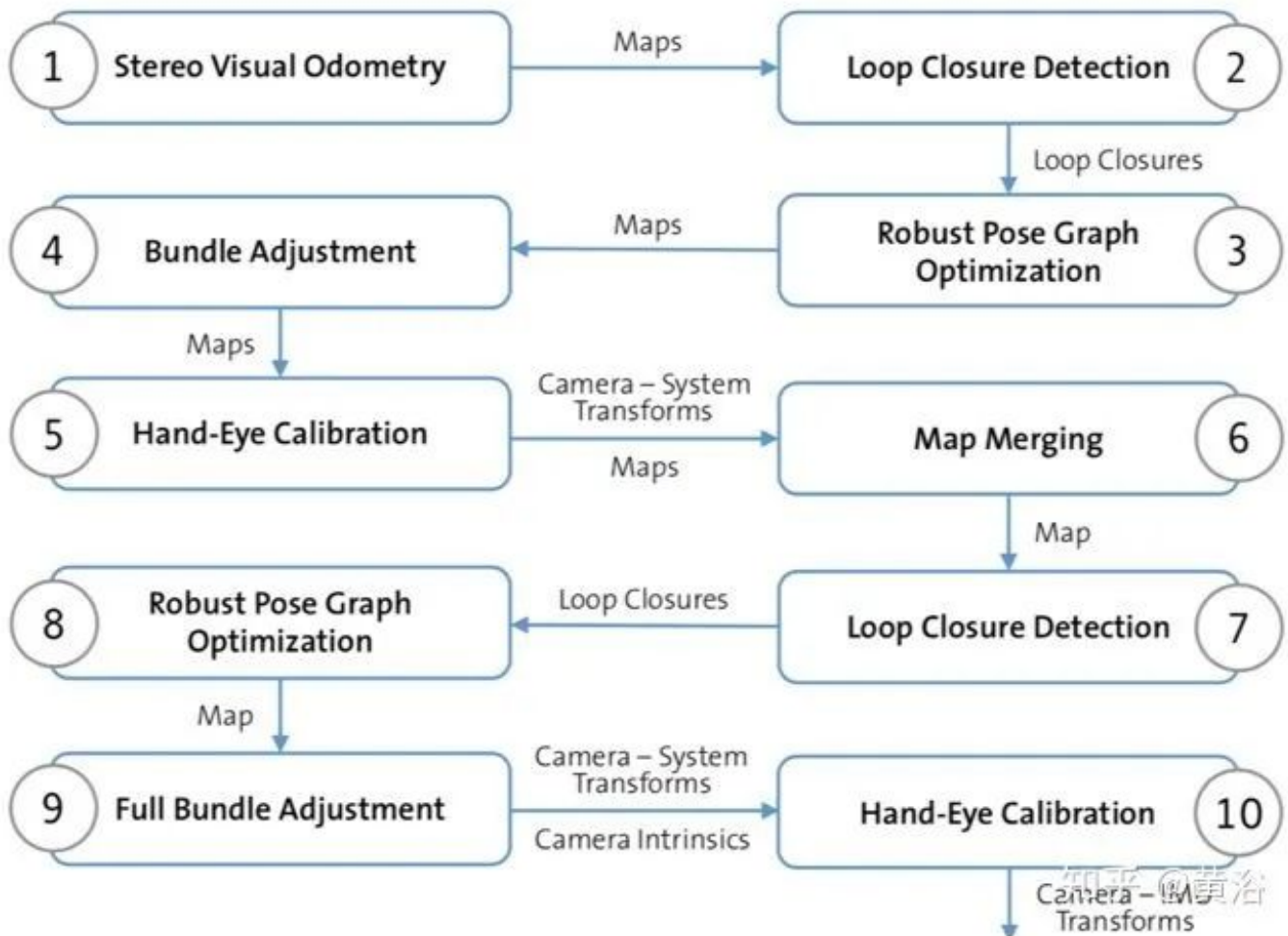
PatchMatch是一个加速图像模版匹配的算法，被用在光流计算和视差估计上。之前微软研究院曾经做过一个基于单目手机相机3-D重建的项目，仿造以前成功的基于RGB-D算法KinectFusion，名字也类似MonoFusion，其中深度图估计就是采用一个修正的PatchMatch方法。

再说在线标定。

这是一个利用路上标志线（斑马线）的标定方法：已知斑马线的平行线模式，计算Homography将模式和路面实现匹配。



这个流程图比较复杂，采用SLAM做在线标定，不适合高频率操作：



和单目方法类似，采用车道线平行和路平面这个假设可以快速完成在线标定，即消失点（vanishing point）理论：跟初始化的消失点（与线下标定相关）比较可以算出双目外参数的漂移量。



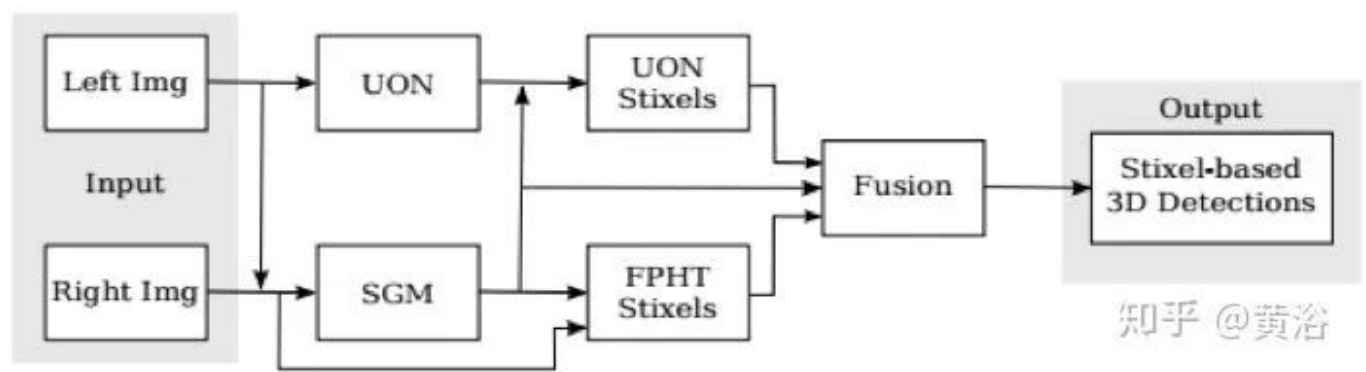
下面介绍几个典型的双目自动驾驶系统。

Berta Benz采用的障碍物检测算法Stixel基于以下假设：场景中的目标描述为列，重心的原因目标是站立在地面上，每个目标上的上部比下部的深度大。下图(a-d)介绍了SGM视差结果如何生成Stixel分割结果：



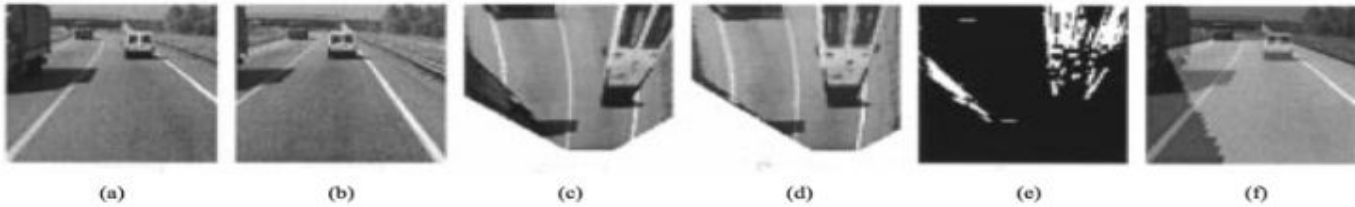


这是他们加上深度学习做视差融合之后再做Stixel的框图和新结果：

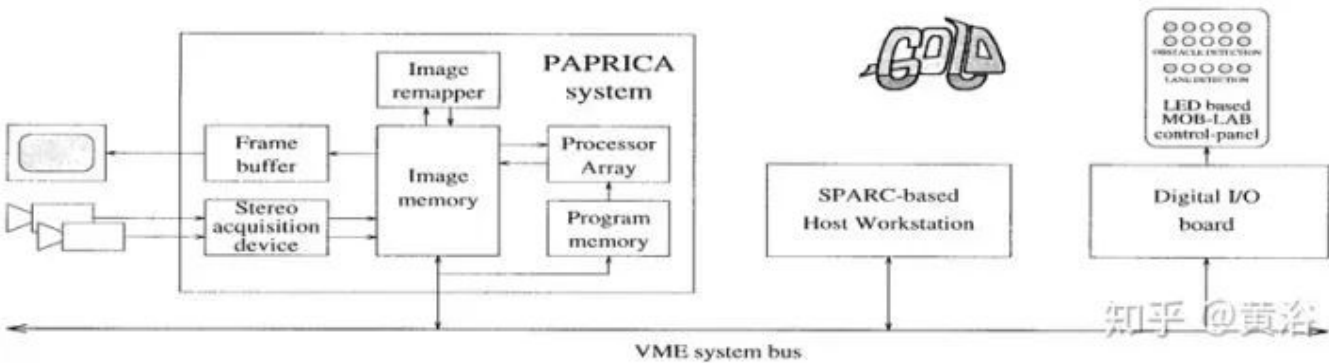
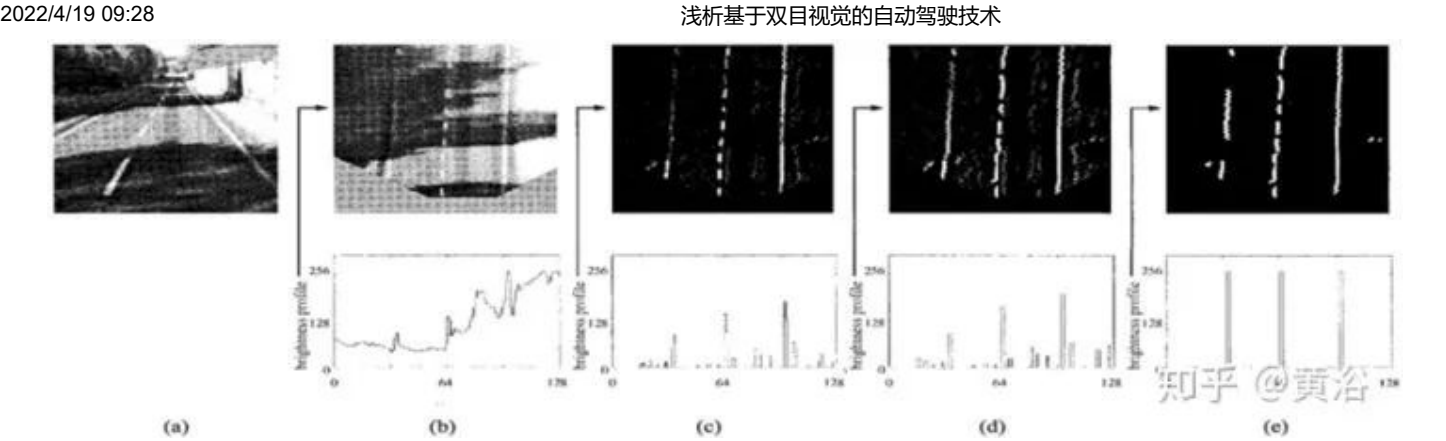




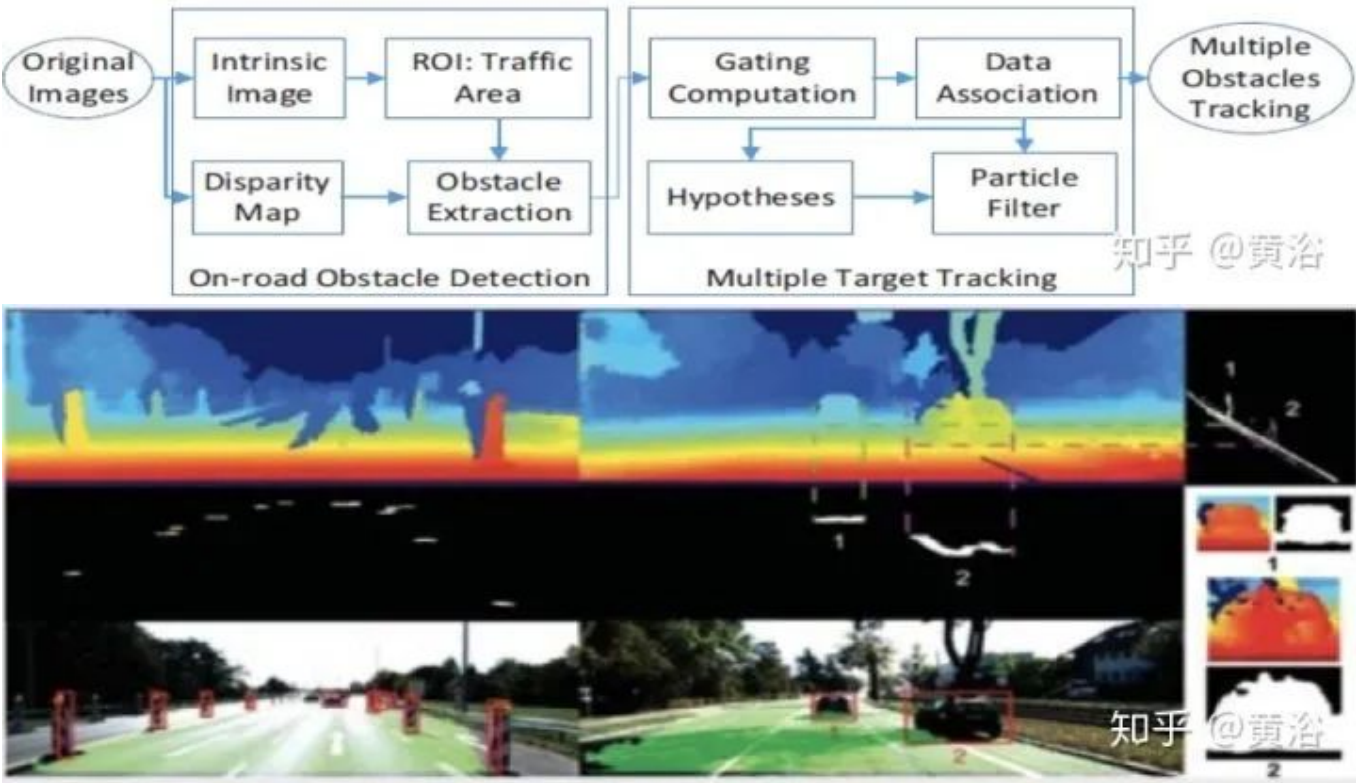
介绍一个早期双目障碍物的算法，Generic Obstacle and Lane Detection system (GOLD)。基于IPM（Inverse Perspective Mapping），检测车道线，根据左右图像的差计算路上障碍物：



(a) Left view. (b) Right view. (c) Remapped left view. (d) Remapped right view. (e) Thresholded and filtered difference between remapped views. (f) In light gray, the road area visible from both cameras.



一种经典的方法是根据路面方程（立体视觉）得到路面视差，基于此计算出路面的障碍物：



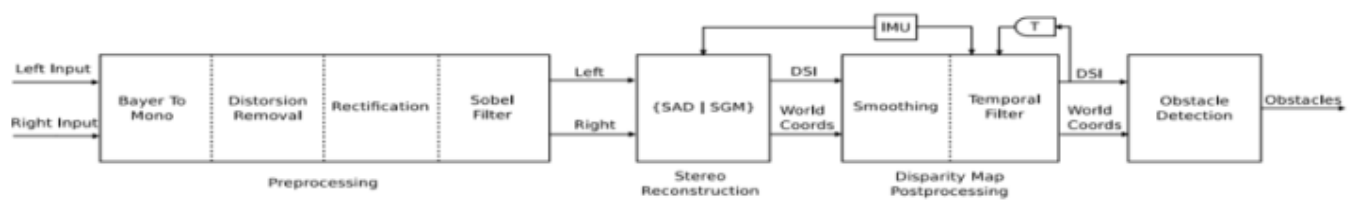
最后介绍一下安霸收购的VisLab双目系统：



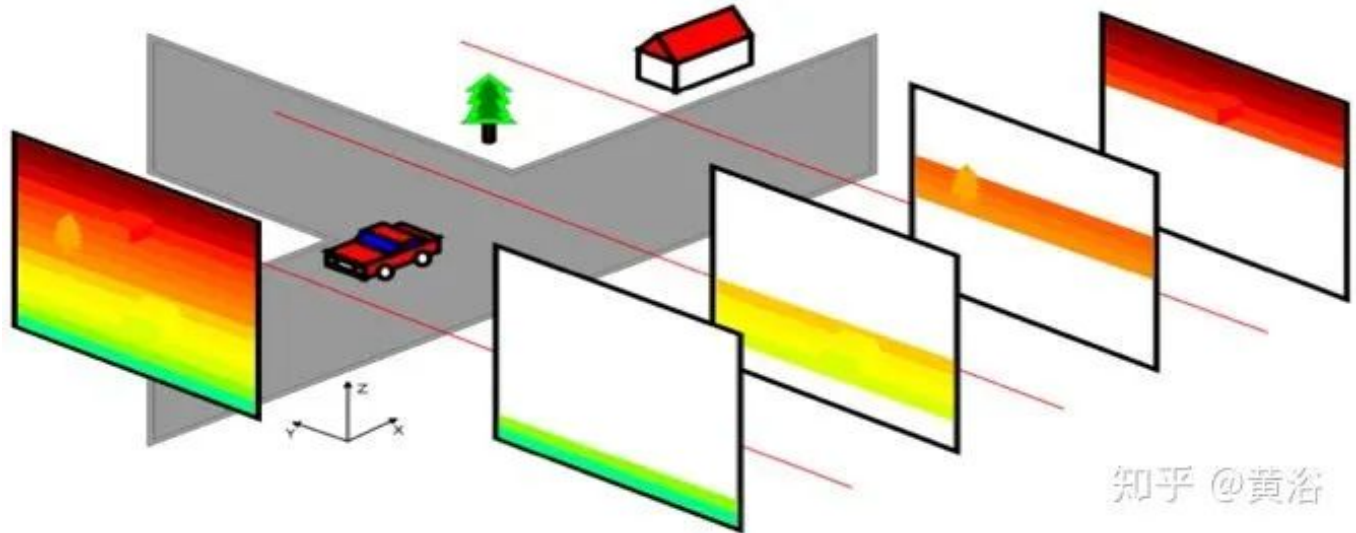


知乎 @黄浴

这是双目障碍物检测流程图：

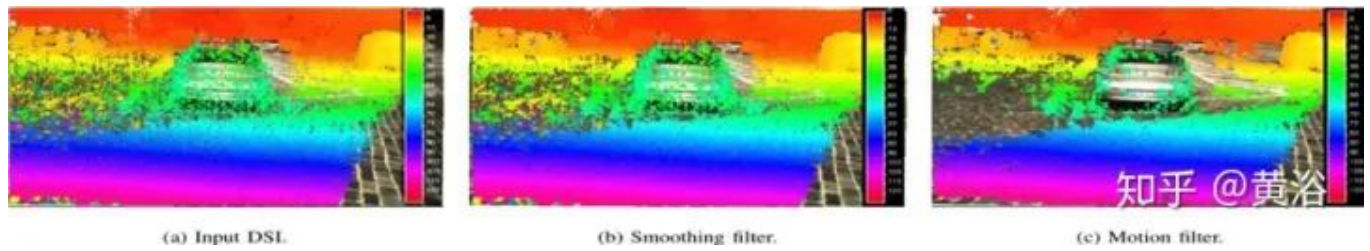


这是基于划分DSI（disparity space image）加速视差估计的方法：



知乎 @黄浴

另外还有提高视差精度的滤波方法：



总的看，双目检测障碍物的方法基本基于视差图，基于路面视差的方法较多。也许随着深度学习发展的突飞猛进，加上计算平台的增强，双目自动驾驶系统也会普及起来。