

StereoBM调研

opencv 官网提供 stereobm 算法原型来自 Dr. Kurt Konolige[1]，下面将介绍该算法原理[2]。

1 立体分析 Stereo Analysis

定义：立体分析是测量目标物体距离相机的一种过程，测量方法基于目标物体在两张或多张图片上的投影[1]。

本质：立体分析的本质问题是找到两张或多张图片中的相关。一旦匹配成功，可利用图像几何（三角形相似原理，成像原理，对极几何等）来计算目标物体到相机之间的距离[1]。

2 匹配方法

匹配算法分类[1]：

- 局部方法：基于图像中的一小片区域特征去匹配另一张图片中的特征区域；
- 全局方法：对局部方法进行补充，考虑物理限制，如表面连续性。

局部方法[1]：根据匹配的是特征是否离散或 correlate a small area patch 分

- Feature-based algorithms
- Area correlation algorithms

Feature-based algorithms[1]：优点是匹配速度较快，能够补偿因视角迁移或更换相机带来的问题，因为场景中的特征点不轻易改变；缺点是测距是稀疏结果，特征点提取很难。

Area correlation algorithms[1]：利用图块 patch 的关联(correlation)关系在不同图片中进行匹配。优点是不计算特征点情况下，可得致密结果；缺点是 area size 选择靠经验，small areas 更有可能得到好的结果，越大的 area 导致高信噪比。

Area Correlation 算法步骤：

- （1）几何矫正 geometry correction：利用 warping 方法将失真的图片变为“标准形式”；
- （2）图片变换 image transform：转换成灰度图，并用平均灰度值（也就是强度值 intensity value）归一化所有像素；
- （3）区域关联 area correlation：实施关联的一步，具体做法是利用一定大小的搜索窗口去搜另一幅图的相似区域；
- （4）极值提取 extram extraction：基于每个像素关联到的最好的两幅图块，产生视差图；
- （5）后期过滤：将视差图中的噪点过滤掉。

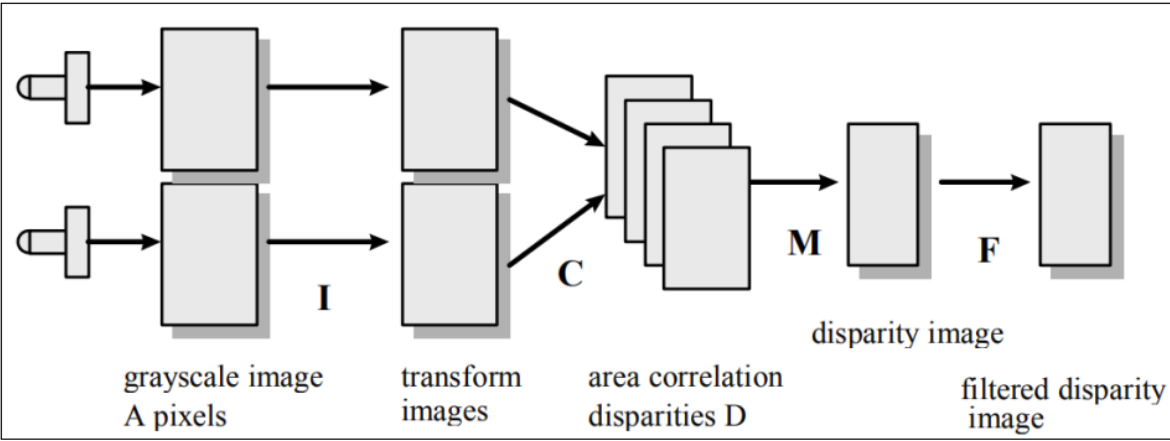


图1 算法流程

3 Correlation和过滤

3.1 Correlation 算法分类

Normalized intensities square difference	Fua 1993
Laplacian of gaussian zero crossings sign square difference absolute difference	Marr and Poggio 1979 Nishihara 1984 Matthies 1995 Kanade 1996
Nonparametric rank census	Zabih and Woodfill 1994

图2 area-based双目 correlation measure

3.2 过滤算法分类

视差图去除噪点

Correlation surface: peak width, peak height, number of peaks	Matthies 1993, Nishihara 1984
Mode filter	
Left/Right check	Fua 1993, Bolles and Woodfill 1993
Interest Operator	Moravec 1984
Interpolation	Nishihara 1984

图3 Range Filtering Operations 测距过滤操作

4 作者的stereobm

作者实践的是在 correlation 和 Posterfiltering部分应用自己的方法，如以下：

- LOG transform, L1 norm (absolute difference) correlation;
- Variable disparity search, e.g., 16, 24, or 32 pixels;
- Postfiltering with an interest operator and left/right check;
- x4 range interpolation

视差计算原理如下图

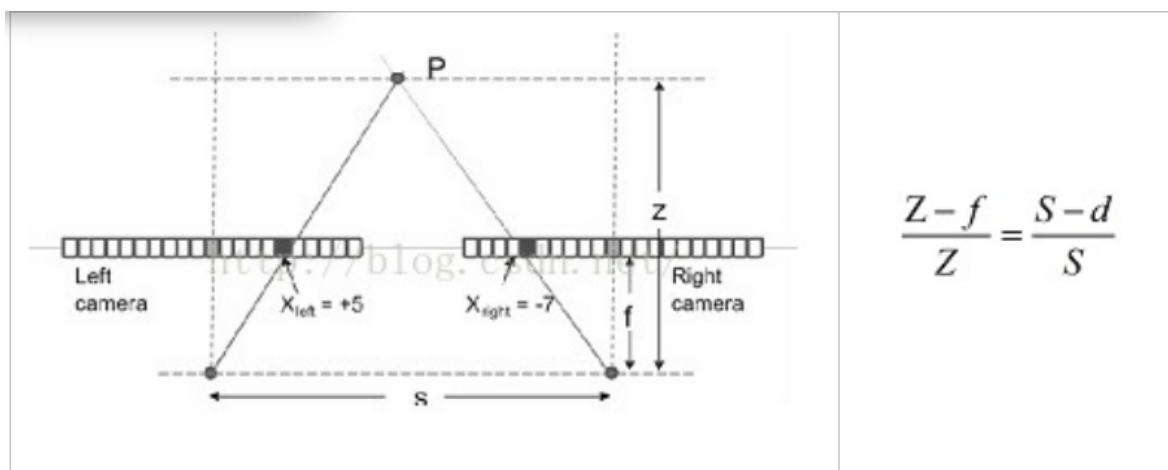


图4 相似三角形原理计算视差

采取两个同一水平线上的摄像头进行拍摄的时候，同一物体将在两个摄像机内被拍摄到，在两个摄像机内部，这个物体相对于摄像机中心点位置有不同的坐标，如上图4所示。Xleft 是该物体在左摄像机内相对位置，Xright是该物体在右摄像机内相对位置。两个摄像机相距 S，焦距为 f，物体P距离摄像机 z，也就是景深。当我们将两幅图像重叠在一起的时候，左摄像机上P的投影和右摄像机上P的投影位置有一个距离|Xleft|+|Xright|，这个距离称为视差 Disparity，根据相似三角形图3可以得到 $z=sf/d$ 。也就是只要计算得到了d 的值，就可以计算得到深度图。而在计算d 的值的过程中需要对两幅图像进行匹配，寻找到物体P在两幅图像中的相对位置。在对图像进行匹配的过程中，需要用到cost computation，即通过寻找同一水平线上两幅图像上的点的最小误差来确定这两个点是否是同一个物体所成的像。由于一个点所能提供的信息太少，因此往往需要扩大对比的范围。

5 opencv 源码解析

[a] <https://blog.csdn.net/ultimate1212/article/details/86079153>

[b] SAD 窗口Matching : <https://blog.csdn.net/KYJL888/article/details/79250481>

[c] <https://zhuanlan.zhihu.com/p/50801189>

参考文献

[1] https://docs.opencv.org/3.4.1/d9/dba/classcv_1_1StereoBM.html

[2] https://www.cs.cmu.edu/~motionplanning/papers/sbp_papers/integrated1/konolidge_stereo_vision.pdf