**基于半全局匹配和互信息的立体图像处理**

摘要：这篇文章介绍了一种半全局立体匹配方法。它使用一种基于互信息的逐像素匹配代价来补偿输入图像的辐射测量差异。逐像素匹配是基于平滑约束进行的，它通常表示为全局代价函数。SGM方法通过各方向路径优化来进行快速近似。本文也解决了遮挡检测，亚像素精确化，以及多基线匹配等问题。另外，本文提出了针对移除轮廓，从结构化环境的特定问题中恢复，以及间断点插值的后续处理步骤。最后，通过标准立体图像对的对比显示出SGM是目前排名最高的算法，而且如果考虑亚像素准确度的话它是最好的。复杂度和像素数量以及视察范围线性相关，对于典型测试图像仅需1-2秒即可得出结果。在基于互信息匹配代价的深度估计中证明了对于大范围的辐射测量变换有一定鲁棒性。最后，进行了几个实验显示出本文的方法在实际问题中应用的很好。摘要：这篇文章介绍了一种半全局立体匹配方法。它使用一种基于互信息的逐像素匹配代价来补偿输入图像的辐射测量差异。逐像素匹配是基于平滑约束进行的，它通常表示为全局代价函数。SGM方法通过各方向路径优化来进行快速近似。本文也解决了遮挡检测，亚像素精确化，以及多基线匹配等问题。另外，本文提出了针对移除轮廓，从结构化环境的特定问题中恢复，以及间断点插值的后续处理步骤。最后，通过标准立体图像对的对比显示出SGM是目前排名最高的算法，而且如果考虑亚像素准确度的话它是最好的。复杂度和像素数量以及视察范围线性相关，对于典型测试图像仅需1-2秒即可得出结果。在基于互信息匹配代价的深度估计中证明了对于大范围的辐射测量变换有一定鲁棒性。最后，进行了几个实验显示出本文的方法在实际问题中应用的很好。

# 简介

精准稠密立体匹配对于许多应用场景如三维重建是一项非常重要的工作。其中的困难之处在于遮挡问题，物体边缘问题，以及结构问题，它通常表现的很模糊。对于无纹理和重复性纹理问题匹配很容易受到挑战，这在结构化环境中表现的很明显。

# 半全局匹配

半全局匹配（SGM）是基于互信息的逐像素匹配和通过组合许多一维约束近似逼近全局二维平滑约束的思想。程序流程图如下，下面依次介绍该算法的几个处理步骤。

|  |
| --- |
|  |
|  |

## 逐像素匹配代价计算

输入的左右图像必须已经知道对极几何模型，要计算参考图（base image）某点的匹配代价，需要用到其灰度为，及在待匹配图（match image）的疑似匹配点，其灰度为，和之间有极线方程，是极线的参数。如果图像已经被校正了，那么待匹配图在参考图的右边，且代表视差。

一个重要方面是考虑匹配区域的大小和形状。 区域面积越大，匹配的鲁棒性越高。但是区域面积变大后以后会导致物体的边界模糊。这里不使用邻域内的视差是连续的这一假设，也就是说，只有和这两个灰度值被用来计算点的匹配代价，其他的点不考虑。

文章使用了新的方法，即基于互信息的匹配代价，该方法对光照不敏感。互信息是基于熵来计算的。熵是用来表征随机变量的不确定性，不确定性越强那么熵的值越大（最大为1），那么图像的熵其实就代表图像的信息量。互信息（MI）度量的是两个随机变量之间的相关性，相关性越大，那么互信息就越大。则立体匹配的问题变为求互信息最大化的问题。下面给出熵和联合熵的定义：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

其中代表某个点的概率分布，即灰度直方图中灰度为的点出现的概率。根据Kim等人的结果将联合熵公式进行泰勒展开可得：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

即各像素点对应的h值之和。计算如下：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

其中是高斯核函数，起平滑滤波作用。 是两幅图像的联合概率密度函数，即归一化之后的统计直方图。计算公式如下式所示：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

其中，指的是像素灰度值对，而不是像素坐标，联合概率密度求完后是一副的图像。同样的，对于熵的计算可用泰勒公式近似简化为下式：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |
|  |  | (3- ) |

最后，基于互信息的定义式可改写为：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |
|  |  | (3- ) |

相应的，基于互信息的某点p的匹配代价为：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

上式表示在计算某一点的匹配代价时，先根据极线方程求出在右图像中对应点的灰度，再根据这两点灰度值构成的灰度值对直接在互信息图M中查找，最后取负值。为了加快算法运行的速度，采用分层互信息(HMI)的思想，即递归的使用降采样过的参考图和待匹配图，经过HMI的计算，SGM算法对光照有了较强的鲁棒性，同时提高了后续迭代求精等操作的效率。

## 代价聚合

构造全局能量函数

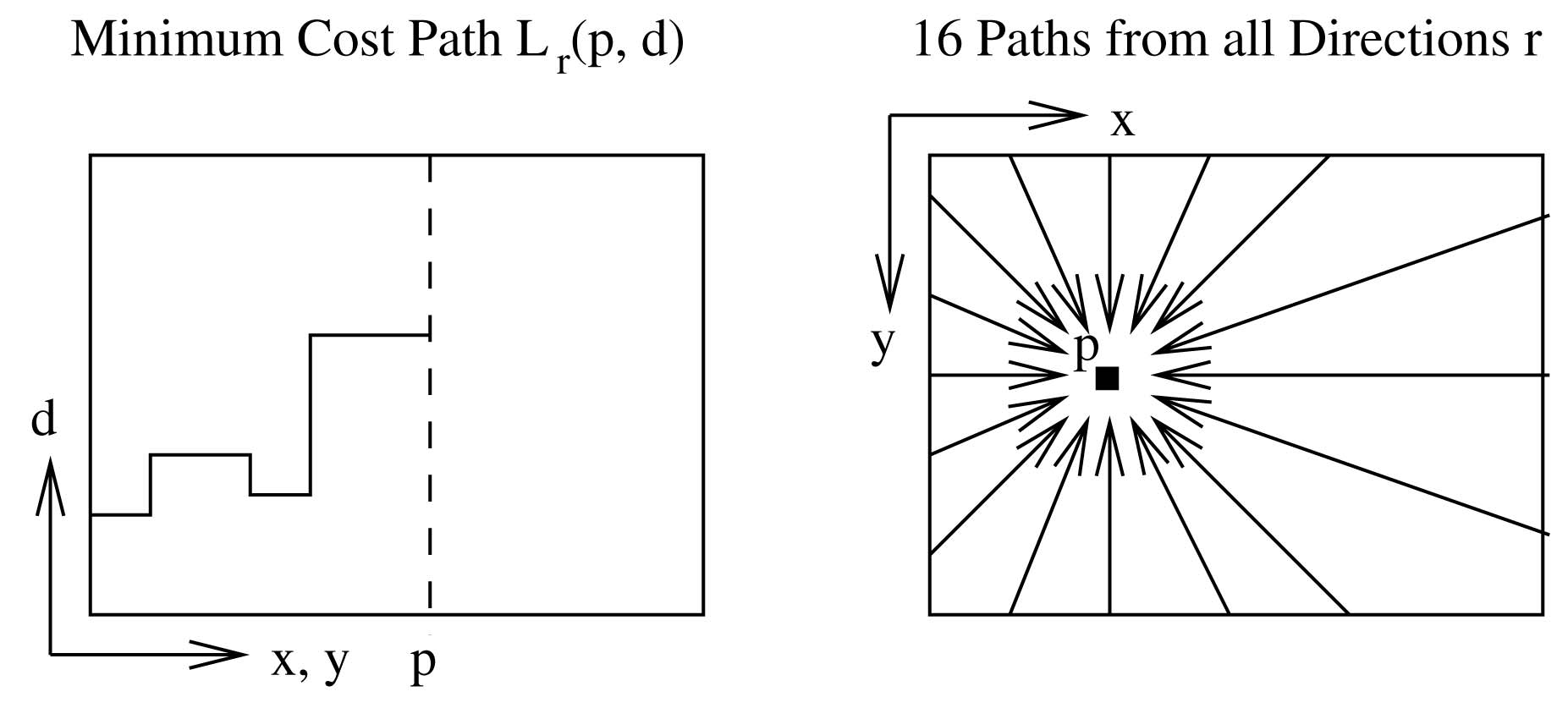
由3.1.1已经得到基于互信息的匹配代价，但逐像素匹配代价容易受到误匹配和噪声点的影响，因此还要考虑使用某点的邻域视差数据来构造惩罚函数以增加平滑性约束。能量函数的定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

上式中代表的是基于互信息的代价计算项，后面两项是指当前像素及其邻域内所有像素之间的约束。如果和之间视差差1，则惩罚；否则惩罚。

动态规划求解E

对于的动态规划求解其实是一个NP-hard问题（non-deterministic polynomial），若在2D图像上进行求解几乎是不可能的。通常的解决办法是在多个一维方向上进行动态规划，然后近似组成二维的平滑约束。因此，我们考虑沿着像素点一圈8个或16个方向进行优化。它使用某点上邻域内所有方向上的代价构成一维的DP问题。



代表某一点的像素聚合值，其计算方法如下：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

以为中心，可以找到16个邻域方向，在每个邻域方向上有一个指向的矢量，表示16个矢量中的某一个。代表当前矢量上的代价聚合值，其计算方法如下：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

点某个邻域方向的代价聚合值分为三项，第一项是当前点的匹配代价，第二项是周围点的代价聚合值的最小值，第三项是为了防止计算结果过大做的统一调整。OpenCV中参数规定如下：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |
|  |  | (3- ) |

是图像的通道数，是初始设定的SAD窗口大小，数值为奇数。

## 视差计算

求解上述，每个点的视差即为对应于最小代价的。

## 后续处理

唯一性检测

检测在匹配图像中的特征点是否是唯一点。在OpenCV中定义一个uniquenessRatio值，当搜索窗口内的最低代价是次低代价的(1+uniquenessRatio/100)倍时，最低代价对应的视差才是该点的视差，否则该点视差为0。

亚像素差值

上述步骤计算出的视差为整数像素，如果直接使用会在后续步骤中产生很大的误差，因此有必要对像素进行细化得到更加精细的视差，即在一个范围内拟合一条曲线，求它在这个范围内的最小值。常见的拟合曲线是抛物线，如下图所示：



从上图中我们可以确定三个点对要从这三个点拟合出一条抛物线，设抛物线的方程为：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

这三个点均在抛物线上，有：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

将三个点带入抛物线方程中，得：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

为简化计算，假设把坐标系原点移到处，即，系数可由克拉默法则得出：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

对于抛物线，令其一阶导数为0，可得：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

将(3- 21)代入(3- 22)中，再将坐标系左移，得：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

而在OpenCV中，亚像素精确化用以下公式实现：

|  |  |  |
| --- | --- | --- |
|  |  | (3- ) |

左右一致性检测(LRC)

在双目视觉中，由于两个相机间隔一定距离，经常会出现某个相机部分被遮挡，在这种情况下匹配效果会变得很差。LRC是用来检测图像的遮挡区域的一种方法，它的执行步骤如下：

1. 给定两个相机的输入图像和,分别以其为基准图像求得视差图和。
2. 对于左图像中某个像素，取。
3. 计算它在右图中的视差，。
4. 若，则对应位置有正确的视差，否则是遮挡点。

连通区域检测

对LRC的结果再次检测误匹配点，统计当前处理的视差点周围满足连通条件的点的个数，若个数小于某个阈值就认为是误匹配点。此检测有助于去除经过唯一性检测和LRC检测后残余的噪声点，效果较为理想。