

# 未定标双摄像头的目标距离估计与自动定位系统的初步实现

朱恬骅

09300240004 计算机科学与技术

2012年6月

- 1 简介
- 2 SIFT和SURF算法概述
  - SIFT算法概述
  - SURF对SIFT的改进
- 3 基于频率域残差的显著性检测
  - 基于频率域残差的显著性检测
  - 图像粗分割
- 4 距离估计方法
  - 定标状态下的距离估计
  - 未定标状态下的距离估计
  - 摄像头共平面共X轴情况下的距离估计
  - 摄像头共面情况下的讨论
- 5 实验
  - 对视频应用SIFT进行对象识别的实验
  - 对视频进行距离估计的实验
  - 与SGBM算法的比较

## 背景

在机器人领域中，如何自动查找兴趣目标，并对目标进行识别和定位，始终是一项重要的工作。由于我们不可能通过超声波定位等方法达到较高的分辨率，也较难将其它非光学定位系统的定位结果与视频图像相结合以达到直观的目的，运用双摄像头进行机器人立体视觉成为了必然的选择。

在传统的立体视觉实现上，通常使用[6]中提出的定标方法，事先对一已知目标进行多视角的采样，以此估计出两个摄像头的内外参数，以建立摄像图像坐标与世界坐标的换算关系，从而对物体的距离进行估计和定位。

较高精度，较多人工参与。

# 动机

- 在一些场景中，我们对物体的距离只需要一个大致的估计，并不需要太高的精度；同时，摄像头的位置由于种种原因可能发生变化。
- 一个“即插即用”的系统。
- 同时，我们希望机器人能够自动寻找感兴趣的对象进行再认。关于对象的具体知识可由其它途径获得，例如可以想像一个网络上的图像索引的知识数据库。

## 相关工作

- ① SIFT/SURF [4] [1]
- ② 基于频域残差的显著性检测[3]

## 方法

- ① 使用SURF匹配两幅图像中的关键点
- ② 使用显著性检测查找可能的目标和细节丰富的背景区域，缩小SURF匹配的空间
- ③ 使用SURF匹配的结果估计距离
- ④ 利用第一步中得到的SURF特征与已保存的对象特征匹配以识别

## 1 简介

## 2 SIFT和SURF算法概述

- SIFT算法概述
- SURF对SIFT的改进

## 3 基于频率域残差的显著性检测

- 基于频率域残差的显著性检测
- 图像粗分割

## 4 距离估计方法

- 定标状态下的距离估计
- 未定标状态下的距离估计
- 摄像头共平面共X轴情况下的距离估计
- 摄像头共面情况下的讨论

## 5 实验

- 对视频应用SIFT进行对象识别的实验
- 对视频进行距离估计的实验
- 与SGBM算法的比较

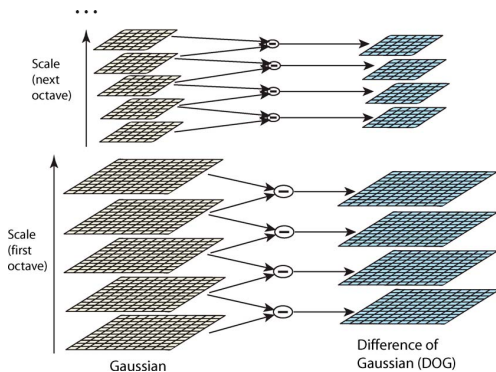
# SIFT算法

- Lowe [4] 2004年总结完善。
- 提取局部特征的算法，在尺度空间寻找极值点，提取位置，尺度，旋转不变量。

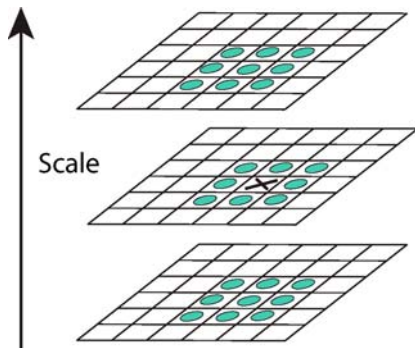


## 尺度空间

尺度空间理论目的是模拟图像数据的多尺度特征。高斯差分尺度空间（DoG scale-space）。利用不同尺度的高斯差分核与图像卷积生成。采用图像金字塔的方法解决这一问题。



## DoG空间极值检测



获得位置尺度信息。

## 确定关键点主方向

为达到旋转不变性，需要根据特征点的局部图像特征找一个方向的基准。

以关键点为中心，用直方图统计一定领域内像素的梯度方向。定义直方图的最大值为关键点的主方向。

## SURF对SIFT的改进

- Speeded-up Robust Features (SURF), 由H. Bay等人在2008年提出。[1]
- SIFT在计算图像金字塔的时候花去了大量的时间, 主要是不断地重复计算高斯卷积。
- SURF用部分和的思想加快了卷积的过程, 称为“积分图像”(integral images)。
- 改变高斯滤波器的大小, 而不是调整图像大小, 来加快图像金字塔的计算。
- 减少使用的维数。Lowe曾推荐使用128维的描述向量, SURF可以只用64维。

## 对高斯滤波器进行近似

H. Bay等人使用盒状滤波器（box filters）来模拟高斯滤波器，称为高斯滤波器的拉普拉斯近似（Laplacian of Gaussian Approximation）。

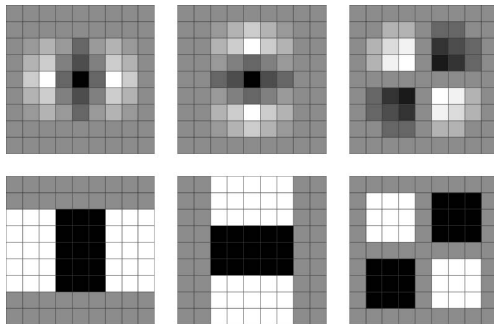


Figure: Laplacian of Gaussian Approximation. [1]

- 1 简介
- 2 SIFT和SURF算法概述
  - SIFT算法概述
  - SURF对SIFT的改进
- 3 基于频率域残差的显著性检测
  - 基于频率域残差的显著性检测
  - 图像粗分割
- 4 距离估计方法
  - 定标状态下的距离估计
  - 未定标状态下的距离估计
  - 摄像头共平面共X轴情况下的距离估计
  - 摄像头共面情况下的讨论
- 5 实验
  - 对视频应用SIFT进行对象识别的实验
  - 对视频进行距离估计的实验
  - 与SGBM算法的比较

## 基于频率域残差的显著性检测

Saliency Detection: A Spectral Residual Approach [3] 2007年提出



## 图像粗分割

- 显著性检测帮助我们去掉了图像中细节不丰富的部分。这些部分对于SURF算法而言也是同样没有意义的。
- 由此可以缩小进行SURF匹配的矩形区域，以加快图像匹配过程的运行时间。
- 假设细节丰富的部分较为集中。之所以不能过于细分，是因为我们需要保留一部分背景信息以供距离估计时参照。
- 简单地对X-Y轴进行记数投影，即可确定一个矩形区域。



- 1 简介
- 2 SIFT和SURF算法概述
  - SIFT算法概述
  - SURF对SIFT的改进
- 3 基于频率域残差的显著性检测
  - 基于频率域残差的显著性检测
  - 图像粗分割
- 4 距离估计方法
  - 定标状态下的距离估计
  - 未定标状态下的距离估计
  - 摄像头共平面共X轴情况下的距离估计
  - 摄像头共面情况下的讨论
- 5 实验
  - 对视频应用SIFT进行对象识别的实验
  - 对视频进行距离估计的实验
  - 与SGBM算法的比较

## 定标状态下的摄像头

- 摄像头定标的目的是建立摄像头坐标与世界坐标之间的转换关系。
- 以[5] 中给出的摄像头成像模型为例：

$$u = S_1 G_1 A T u_0 \quad (1)$$

其中  $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$  满足  $\det(A) > 0$ ,  $u_0$  为现实世界中的图像。

# 摄像头模型

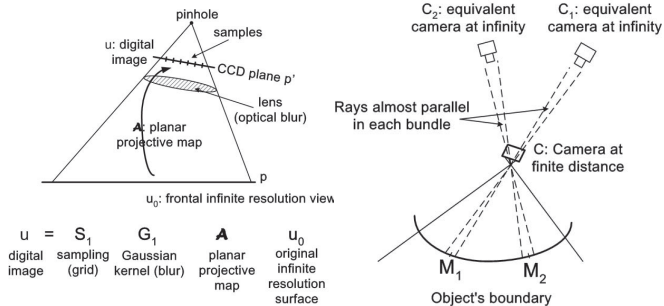
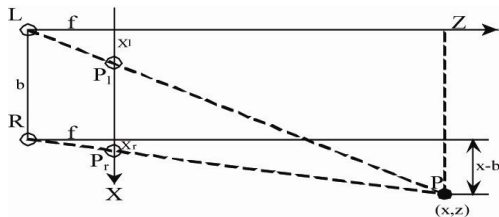


Figure: 摄像头模型[5]

## 未定标状态下的距离估计

- 当摄像头未定标时，我们无法建立“真实的”摄像头坐标与世界坐标的转换关系，只能做到摄像头之间的坐标转换。
- 由于不同摄像头之间内部参数不同（如焦距、扭曲变形情况等），为简化问题，我们采用相同型号的两个摄像头，并忽略它们的内部参数差异。
- 摄像头之间的坐标转换问题简化为了因摄像头位置不同所带来的误差问题。

## 摄像头共平面共X轴情况下的距离估计



目标点 $P$ 在左右图像中的坐标分别为 $(x_l, y_l)$ 和 $(x_r, y_r)$ 。由于共X轴， $y_l = x_l$ ，所以有距离 $z$ 满足：

$$z = \frac{fb}{x_l - x_r} = \frac{fb}{x_l - x_r} \quad (2)$$

我们可以很方便地将其扩展到摄像头共平面X轴互相平行的情况。

## 摄像头共面的情况

摄像头共面的情况下，我们将引入一个新的参数 $\theta$ ，表示两个摄像头之间X轴的夹角。这将复杂化我们的问题，因为我们的摄像头没有图像用于事先的定标以确定这一夹角。我们也无法保证图像中一定有含直线的图形，且这一图形的角点会被SURF捕捉到（因SURF对边界响应比较高，它的边界点通常会被忽略）。对此，我们可以考虑采用Hough变换，在图像中寻找出一条或多条直线，通过比较这些直线之间的夹角来估计 $\theta$ 。

# Hough变换

Hough变换的作用是在图像中找出可能的直线。其思路是：将一条直线用两个参数表示，分别为直线的倾角 $\alpha$ 和离开原点的距离 $\rho$ 。对于每一对参数 $(\alpha, \rho)$ ，在图像中寻找这一直线上的点在图像中为边缘的数量。点数越多的，则这一图像中含有下列方程所决定的直线的可能性越大：

$$x \cos(\alpha - \pi/2) + y \sin(\alpha - \pi/2) = \rho. \quad (3)$$

我们可以认为： $\theta \simeq \text{mean}(\alpha_l - \alpha_r)$ 。

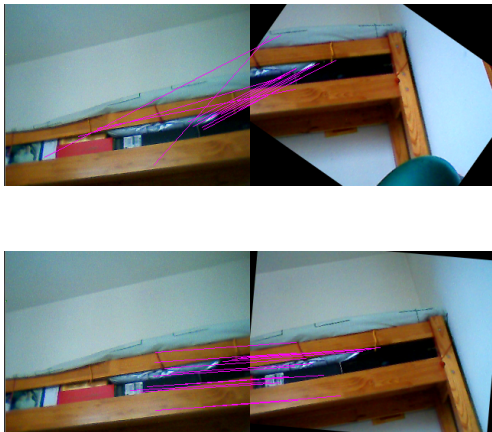
如果我们引入一个表示尺度的因子 $\lambda$ ，我们就可以解决平行平面内两个摄像头的坐标转换问题了。可以使用同一级别距离上的特征点之间的对应关系计算得到，此处从略。

## 确定图像主方向

- 理想状态下，我们应当确定左右图像中直线的对应关系，以获得较为准确的 $\theta$ 。这可以通过使用SURF得到。
- 在本项目中，我们采用的是类似SIFT中确定主方向的思想：对每一幅图像中的倾角 $\alpha$ 用直方图统计，选取其中数量最大的一个方向为图像的主方向。
- 定义两个主方向的差就是所要估计的 $\theta$ ，其精确度由直方图的间距决定。

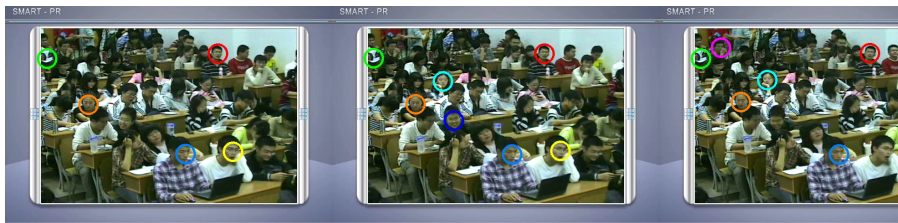


## 旋转矫正

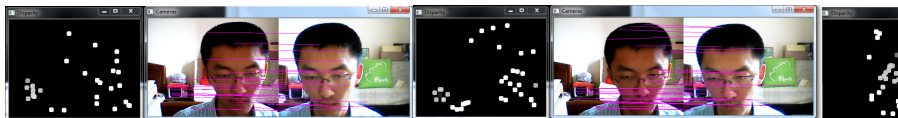


- 1 简介
- 2 SIFT和SURF算法概述
  - SIFT算法概述
  - SURF对SIFT的改进
- 3 基于频率域残差的显著性检测
  - 基于频率域残差的显著性检测
  - 图像粗分割
- 4 距离估计方法
  - 定标状态下的距离估计
  - 未定标状态下的距离估计
  - 摄像头共平面共X轴情况下的距离估计
  - 摄像头共面情况下的讨论
- 5 实验
  - 对视频应用SIFT进行对象识别的实验
  - 对视频进行距离估计的实验
  - 与SGBM算法的比较

# 对视频应用SIFT进行对象识别



## 对视频进行距离估计



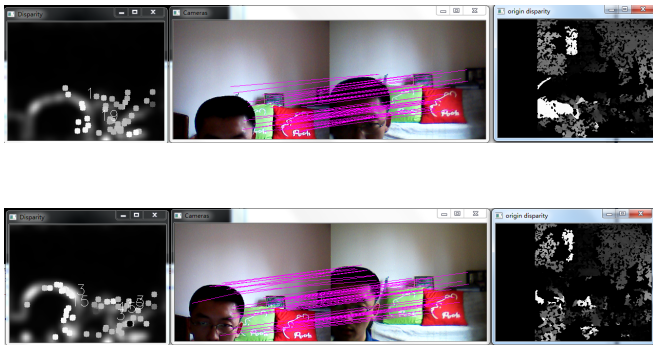
## 与SGBM算法的比较

Semi Global Block Matching (SGBM) [2] 是OpenCV 2.3 中的新特性，提供了一个同样是非标定情况下对象距离估测的算法。

想法上有两点不同：

- ① SGBM考虑的范围是整幅图像，本项目中考虑的是特征点的距离。
- ② SGBM的结果需要进一步处理才能用于对象识别，我们可以用特征点直接参与对象识别。

## 与SGBM算法比较的结果





Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool.

Speeded-up robust features (surf).

*Comput. Vis. Image Underst.*, 110(3):346–359, June 2008.



Heiko Hirschmuller.

Accurate and efficient stereo processing by semi-global matching and mutual information.

*In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02*, CVPR '05, pages 807–814, Washington, DC, USA, 2005. IEEE Computer Society.



Xiaodi Hou and Liqing Zhang.

Saliency detection: A spectral residual approach.

*In CVPR*, 2007.