# RGB图像/视频中基于模型的3D物体跟踪问题描述

## 引言

本文参照PWP3D算法 [1]提供的框架思路，阐述：给定一帧RGB图像（或多帧连续图像组成的视频），其场景中包含了已知几何物体，另给定此几何物体的CAD模型，如何通过基于梯度的非线性优化方法，求解模型在图像所呈现的场景中的正确3D姿态{R, t}问题。PWP3D算法[1]将前景（即，目标物体）的轮廓表述为SDF（signed distance function）嵌入函数的零水平集，并基于逐像素后验概率，在前景以及背景区域上建立能量函数E；通过将模型姿态 {R, t} 化为参数 (t1, t2, t3, qw, qx, qy, qz) 7个参数，并表示出函数E在各参数上的偏导，进而通过梯度下降等方法，直接优化模型姿态，使得前景和背景的统计模型差异最大化。

## 符号表示

下表阐述了本文档中所出现的符号，以及其释义。

表 1 文档中出现的符号表示及其释义

|  |  |
| --- | --- |
| 符号 | 含义备注 |
| I | 图像（Image） |
|  | 图像定义域 |
|  | 前景图像区域 |
|  | 背景图像区域 |
|  | 前景统计概率模型 |
|  | 背景统计概率模型 |
|  | 图像中像素坐标 |
|  | 图像像素值，本文中即RGB值 |
|  | **X**为相机坐标系下的点坐标 |
|  | **X0**为物体坐标系下的点坐标 |
| **(R, T)** | R为旋转矩阵，T为平移向量 |
|  | 将R表示为四元数后，(R, T)变为7参数，即前述 (t1, t2, t3, qw, qx, qy, qz) |
|  | 图像中物体可视部分的轮廓 |
|  | 以C为零水平集的水平集函数 |
|  | 即u、v方向的焦距，量纲：像素 |
|  | 平滑赫维赛德阶跃函数 (smoothed Heaviside step function) |
|  | 平滑狄拉克函数(smoothed Dirac delta function) |
| SDF | signed distance function |

## 问题描述

给定图像I，三维CAD模型，求解：模型在图像I中的正确姿态参数，采用的是非线性迭代优化方法。对于每一帧，其第k次迭代，分步的求解目标为:

(a) 使用当前姿态值，对模型进行2D投影；

(b) 模型投影轮廓；

(c) 由轮廓求SDF，即；并计算SDF对x, y的偏导（[1]公式(12)用到）；

(d) 由图像I，前景分割mask，生成前景、背景的统计模型

(e) 由(c)(d)结果为输入，计算能量函数E对待求参数的偏导；

(f) 梯度下降(GD/CG或GD/CG-LS) 方法求解

算法流程图如图1所示。



图 1 PWP3D算法流程

【注】任光阔[2]的测试报告中，提到修改算法流程，求解能量函数E的值，以期之后再E上附加惩罚项。

以下小节3.1~…形式化表达上述(a)~(f)分布步骤中的一些优化目标。

### 水平集函数

对图像中闭合轮廓的描述，采用水平集函数这样的隐式表达，以简化轮廓的描述方式。采用有符号距离函数SDF形式的水平集函数，其定义为：

(1)

### 能量函数表示

PWP3D算法[1]采用基于区域的分割技术，求解能使前景、背景区域的数值统计量差异最大化的闭合轮廓线，其能量函数形式化表达如下：

(2)

[3]将(2)式重写为水平集形式的基于区域的分割：

(3)

其中，由像素属性似然值给出，即 ，其中M为之一。

假设像素独立性，并将(3)式的积分化为离散形式，给定一帧图像可得：

(4)

进而得到”-log”表示形式的后验概率：

(5)

(6)

其中，定义如下：

(7)

(8)

而分别为前景、背景区域面积：

(9)

最后，根据[4]所述，赫维赛德阶跃函数的平滑近似为：

(a)

### 能量函数对姿态参数的微分

如论文[1]公式(10)~(14)所示，基本原理即复合函数求导，略过。

ps：另外[1]中讨论的优化算法选取(4.1节), 多目标物(4.2节), 多视角镜头(4.3节), 不可见点(4.4节), 时序一致性约束(4.5节), 暂时也不在本报告描述范围内。

## 参考文献

[1] Prisacariu V A, Reid I D. PWP3D: Real-time segmentation and tracking of 3D objects[J]. International journal of computer vision, 2012, 98(3): 335-354.

[2] 任光阔, 《分割实验测试报告\_思路整理》

[3] Cremers D, Rousson M, Deriche R. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape[J]. International journal of computer vision, 2007, 72(2): 195-215.、

[4] <https://en.wikipedia.org/wiki/Heaviside_step_function>

## 附录A. PWP3D[1]阅读遗留问题

### He & Phi 复合函数表示是否矛盾？

以公式(9)为例：

按照论文公式(1)对定义，图像上的SDF值，以闭合轮廓线为参照，为**内负外正**（[1]中Figure 2佐证）；而对于He函数，其值以Y轴为界，**左0右1**。两函数复合之后结果，应为是背景面积，是前景面积，与[1]中文本表述相反，但是此前公式(3, 4, 6)采用同样描述，所以是否我理解反了？（比如其实现的He与[4]的定义，公式(a)相悖?）

应结合代码实现去理解，【未解决】 //2016-5-2 23:51:29

### [1]中 table 2 的GD、GD-LS 谁是 with line search?

[1]提到:

*using gradient descent with (GD) and without (GD-LS) a line search step*

但是，按照table 2, GD-LS 迭代次数更少，说明其实 GD-LS 才是用(with)线搜索的方式? ([1]第6页佐证)

【未解决】 //2016-5-4 00:04:56