

基于三维校正的SHGC物体三维结构提取方法

吴志红^{1,2}, 刘日晨^{1,2}, 甘霖³

(1. 四川大学计算机学院 成都 610065; 2. 四川大学视觉合成重点学科实验室 成都 610065; 3. 四川电力科学研究所 成都 610072)

【摘要】针对SHGC(straight homogeneous generalized cylinders)图像的特点与从单幅图像中提取三维信息困难,设计了一种基于单幅SHGC图像和三维校正的三维结构提取方法。与传统的SHGC三维结构提取方法不同,该方法从三维空间中匹配的角度出发,构造一个二维横截面,在二维横截面上增加一个z轴,通过平移、旋转预先构造好的三维横截面,将构造好的横截面与原图像中SHGC物体的横截面进行兴趣点匹配,从而确定最优化结果时的仿射成像参数 (λ, θ) ,通过成像参数可以成功提取SHGC横截面的三维信息,结合SHGC的子午线轮廓信息定性地得到整个SHGC物体的三维信息,提取SHGC表面的纹理,通过三角网格的方法将纹理贴到三维物体表面。实验论证了本文提出的三维结构提取方法是可行的和正确的。

关键词 仿射变换; SHGC; 单幅图片; 三维

中图分类号 TP391.41

文献标识码 A

doi:10.3969/j.issn.1001-0548.2011.03.023

3D Structure Extraction Method Based on Three-Dimensional Rectification for Straight Homogeneous Generalized Cylinders

WU Zhi-hong^{1,2}, LIU Ri-chen^{1,2}, and GAN Lin³

(1. College of Computer Science, Sichuan University Chengdu 610065;

2. State Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University Chengdu 610065;

3. Sichuan Electric Power Research Institute Chengdu 610072)

Abstract A three-dimensional (3D) structure extraction algorithm based on a rectification method in 3D space for SHGC (straight homogeneous generalized cylinders) is designed in this paper. This algorithm resorts to the properties of SHGC and aims at the difficulties of the 3D information extraction from a single image. Different from the conventional algorithms of SHGC, a two-dimensional cross section which resemble the cross section of the original object is constructed in this algorithm. The cross section is placed into a 3D coordinates system with an additional z-axis, which is perpendicular to the 2D cross section. The interesting points on the constructed cross section are matched with those on the cross section of the object from the image. The affine imaging arguments pair, (λ, θ) under the circumstance of the best matching solution are searched step by step through the rotation and transportation operated on the constructed cross section. With the best result pair (λ, θ) , the 3D information of the cross section can be extracted quickly, and the whole 3D structure can be obtained at the next step if the contour is extracted by the meridian of the SHGC. The texture of the SHGC can be extracted and is pasted onto the 3D structure through a mass of triangular grids. The experiments show that the proposed algorithm is available and proper.

Key words affine transformation; SHGC; single view; three-dimensional

因为成像是一个三维到二维的过程,是一个信息损失的过程,如果没有足够的先验知识,要从单幅图像中提取三维信息是十分困难的。但是,如果能抓住一些特殊物体的特点,也有可能从单幅图像中提取三维信息,如SHGC(straight homogeneous generalized cylinders)。文献[1]提出了SHGC的很多数学性质,如几何不变量等,这些性质被接下来的很多研究人员引用^[2-6]。SHGC物体的表面是通过一

横截面沿着一条对称轴进行缩放形成的,并且在人造物体中广泛存在^[7]。

文献[8]提出了有目的的、定性的主动视觉概念,认为马尔视觉计算理论为代表的理论将视觉规定为由场景的图像精确地获取三维信息的机器视觉,要求太高,且没有必要。因此,可以用定性视觉代替定量视觉提取三维信息。

有很多关于提取SHGC图像三维信息的文献,如

收稿日期: 2010-03-15; 修回日期: 2011-01-12

基金项目: 国家自然科学基金(60832011, 60705005); 国家863探索导向课题(2007AA01Z332); 教育部新世纪人才支持计划(NCET-08-0370); 高等学校博士学科点专项科研基金(20090181110052)

作者简介: 吴志红(1964-), 女, 副教授, 主要从事图像处理方面的研究。

文献[2]与文献[7]中都分别有各自的很好的方法。但是文献[2]只是得到SHGC物体表面的横截面曲线,并没有得到真正意义上的横截面三维信息;而文献[7]中给出了针对SHGC物体表面的子午线轮廓提取,但没有得到横截面的三维信息。

本文设计了一种基于单幅SHGC图像的横截面三维信息提取算法,利用横截面成像的近似仿射变换性质,先建立粗略的三维结构信息,再把建立的三维结构经过旋转和平移校正后投影到二维图像坐标系上,将得到的横截面与原图像的横截面进行匹配,从而达到三维空间中校正的目的,结合文献[7]中SHGC的表面轮廓提取方法,最终得到整个SHGC物体的三维结果,并在三维结构的表面构建网格,将纹理贴到三维结构上,从而达到定性视觉所要求的目标。

文献[1]只是提出了一些SHGC物体的性质,并

$$\mathbf{R}_{3D}(\mathbf{n}, \theta) = (\mathbf{p}' \quad \mathbf{q}' \quad \mathbf{r}') = \begin{pmatrix} n_x^2(1 - \cos \theta) + \cos \theta & n_x n_y(1 - \cos \theta) + n_z \sin \theta & n_x n_z(1 - \cos \theta) - n_y \sin \theta \\ n_x n_y(1 - \cos \theta) - n_z \sin \theta & n_y^2(1 - \cos \theta) + \cos \theta & n_y n_z(1 - \cos \theta) + n_x \sin \theta \\ n_x n_z(1 - \cos \theta) + n_y \sin \theta & n_y n_z(1 - \cos \theta) - n_x \sin \theta & n_z^2(1 - \cos \theta) + \cos \theta \end{pmatrix} \quad (1)$$

式中,向量 \mathbf{p}' 、 \mathbf{q}' 、 \mathbf{r}' 分别表示矩阵 $\mathbf{R}_{3D}(\mathbf{n}, \theta)$ 的3个列向量。

性质1 在三维欧式空间中,从坐标系 $\{x''\}$ 到坐标系 $\{x'''\}$ 只有旋转变换时,若把该变换写成 3×3 旋转矩阵 \mathbf{R} ,则 \mathbf{R} 可以用3个 $\mathbf{R}_{3D}(\mathbf{n}, \theta)$ 的乘积表示,即:

$$\mathbf{R} = \mathbf{R}_{3D}(n_1, \theta_1) \mathbf{R}_{3D}(n_2, \theta_2) \mathbf{R}_{3D}(n_3, \theta_3) \quad (2)$$

证明 因为旋转矩阵 \mathbf{R} 是一个三阶的正交阵,如果它有3个不同的特征向量,则表示可以分别绕3个不同的轴旋转, \mathbf{R} 可以用式(2)等号右端3个分别绕不同轴(特征向量)旋转的旋转矩阵 $\mathbf{R}_{3D}(\mathbf{n}, \theta)$ 相乘来表示,此时式(2)中任意一个 $\mathbf{R}_{3D}(\mathbf{n}, \theta)$ 均是不等于单位矩阵的正交阵。

如果有2个相同的特征向量,即表示只绕两个单位向量 \mathbf{n} 旋转,可以把其中任意一个 $\mathbf{R}_{3D}(\mathbf{n}, \theta)$ 用单位阵表示,此时式(2)可以写成:

$$\mathbf{R} = \mathbf{R}_{3D}(n_1, \theta_1) \mathbf{R}_{3D}(n_2, \theta_2) \quad (3)$$

如果3个特征向量都相同,即表示只绕一个单位向量 \mathbf{n} 旋转,则可以把其中任意两个 $\mathbf{R}_{3D}(\mathbf{n}, \theta)$ 用单位阵表示。此时公式(2)可以写成:

$$\mathbf{R} = \mathbf{R}_{3D}(\mathbf{n}, \theta) \quad (4)$$

性质1在本文提取SHGC三维结果时用于旋转目标物体,构造横截面,以校正摄像机拍摄时产生的误差,从而提高提取的三维结构的精度。

2 三维空间中校正并提取三维结构

本文设计了一种以逆向校正的方法提高提取

未给出怎样得到SHGC物体三维信息的方法,而文献[7]确实定性地重建了SHGC物体,但是其中的方法未考虑横截面重建,没有给出横截面重建的具体方法。文献[9]针对的是SHCC(straight homogeneous circular cylinder)物体的位姿估计,SHCC是SHGC的子集。文献[9]在对SHCC物体的3D信息获取时,使用了横截面二维形状的先验知识,但没有考虑物体的横截面重建。本文提出了一种在三维空间中采用逆向校正的方法提取SHGC物体三维信息,尤其是物体的横截面重建的方法,该方法从三维空间的角度出发进行校正,符合成像从三维空间到二维空间的规律。

1 基本理论与几何性质

已知三维空间(x-y-z)中某向量绕单位向量 \mathbf{n} 旋转 θ 角的旋转矩阵公式为:

$$\mathbf{R}_{3D}(\mathbf{n}, \theta) = (\mathbf{p}' \quad \mathbf{q}' \quad \mathbf{r}') = \begin{pmatrix} n_x^2(1 - \cos \theta) + \cos \theta & n_x n_y(1 - \cos \theta) + n_z \sin \theta & n_x n_z(1 - \cos \theta) - n_y \sin \theta \\ n_x n_y(1 - \cos \theta) - n_z \sin \theta & n_y^2(1 - \cos \theta) + \cos \theta & n_y n_z(1 - \cos \theta) + n_x \sin \theta \\ n_x n_z(1 - \cos \theta) + n_y \sin \theta & n_y n_z(1 - \cos \theta) - n_x \sin \theta & n_z^2(1 - \cos \theta) + \cos \theta \end{pmatrix} \quad (1)$$

SHGC三维结构精度的方法,该方法主要分为以下几个步骤:

1) 如果目标物体的有效焦距足够大,导致径向失真过大,则先径向校正单幅SHGC图像,反之跳到下一步。

使用普通摄像机,摄像机的成像过程中,通常最重要的偏差是径向失真,而且随着镜头焦距减小,该误差会更显著^[10]。

因此,如果径向失真无法容忍,则需径向矫正,可以通过最小化代价函数的方法获取失真参数,或者通过Tsai标定等摄像机标定方法获得失真参数^[11]。获取失真参数后就可以通过失真参数形变插补进行径向校正。

2) 三维空间校正并提取三维结构信息。

提取SHGC三维结构,拍摄图像时,摄像机与SHGC物体之间的位置与方位时有以下两点要求:

① 让SHGC物体的对称轴竖直并尽量通过图像的中心;② 摄像后图像上必须能看到SHGC物体其中的一端横截面。此外,定性视觉允许获取的三维信息与世界坐标系中的三维坐标存在一个缩放比例因子。

因为现实世界中的SHGC物体一般都是对称的,基于该对称性特点,可以先确定该SHGC物体的对称轴,再提取物体的子午线,最终可以计算关于对称轴对称的物体的轮廓信息。因此,对SHGC的横截面恢复是提取信息的一个难点。因为拍摄SHGC图像时

要满足上述所提出的位置与方位要求, 本文把从SHGC横截面的像到SHGC横截面的正视图的变换近似看成是仿射变换, 大大减少了射影变换带来的差异。仿射变换可以用以下矩阵表示为:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (5)$$

或表示为分块形式:

$$\mathbf{x}' = \mathbf{H}_A \mathbf{x} = \begin{pmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{x} \quad (6)$$

式中, 仿射矩阵 \mathbf{A} 是一个非奇异矩阵^[10], 总能通过SVD分解分解为:

$$\mathbf{A} = \mathbf{R}(\theta)\mathbf{R}(-\phi)\mathbf{D}\mathbf{R}(\phi) \quad (7)$$

其中:

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \lambda_2 \end{pmatrix} \quad (8)$$

式中, $\lambda_1\lambda_2 = \det\mathbf{A}$ 。

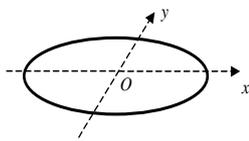


图1 SHGC横截面坐标分解示意图

按照前面所述的摄像机与目标物体的位置与方位约束条件可以知道, 上述仿射变换是在 x 轴与 y 轴两个方向上进行缩放, 再加上定性视觉可以允许存在一个比例缩放因子, 也就是降低了一个自由度(dof), 因此可以不考虑横截面 x 轴方向的缩放因子, 即只要得到 y 轴方向的一个缩放比例 λ' 即可, 此时的 λ' 可以表示为:

$$\lambda' = \frac{\lambda_2}{\lambda_1} \quad (9)$$

式中, λ_1 与 λ_2 分别表示上述仿射变换在 x 轴方向与 y 轴方向的缩放比例因子。

因此, 可以得到提取SHGC物体度量三维信息的步骤:

1) 在SHGC物体表面的子午线轮廓线上提取兴趣点。首先把该轮廓线看成一条曲线, 在曲线斜率大时, 兴趣点可以选择多些。进一步的子午线轮廓提取方法可以参照文献[7]中所述的方法, 即直接计算对称轴两端在 x 轴上的截距。

2) 预处理原图像(图像校正)。在构造横截面之前, 将原二维图像在二维平面内旋转 \mathbf{R} , 平移 \mathbf{T} , 使目标物体的对称轴与图像中心重合。旋转可以按照式(2)~(4)计算。正因为前面的约束条件的约束, \mathbf{R} 一般为单位阵, \mathbf{T} 一般为 $\mathbf{0}$ 矩阵。 \mathbf{R} 和 \mathbf{T} 的具体确定方

法, 可以均匀选取目标物体表面的兴趣点进行计算, 还可以以这些兴趣点为角点进行摄像机标定相对位置信息。本文的图像校正是矫正拍摄过程引起的误差, 而径向校正则是通过摄像机标定消除摄像机内部参数引起的畸变。

3) 构造横截面。横截面三维信息的提取是关键难点, 假设SHGC图像的顶端横截面在 x 轴方向上的截距为 d (单位为像素), 令 $\lambda = 1.0$, 在图像坐标系中构造以下一个横截面 π , 有:

$$\begin{cases} \pi \text{的 } x \text{轴方向上的截距} = d \\ \pi \text{的 } y \text{轴方向上的截距} = d \times \lambda \end{cases} \quad (10)$$

将上述横截面在图1的基础上增加一个 z 轴, 构成三维右手坐标系。利用式(4)将构造的横截面沿 x 轴旋转 θ 角, θ 的初值等于摄像机中心与目标物体中心所成连线与水平线之间的夹角, 平移分量与原图像一致。也正是因为有上述约束条件, 可以想像, θ 比较小, 所以也可以估计 θ , 大部分情况下, $10^\circ \leq \theta \leq 15^\circ$ 。但是该 θ 值只是初值, 还要通过优化得到最优 θ 值。

4) 构造的二维横截面与原SHGC图像的横截面进行匹配。计算所有在横截面上选择的兴趣点的误差平方和, 选择兴趣点如图3所示。用竖直直线将兴趣点两两联系, 形成兴趣点对, 且兴趣点的选择密度与曲线斜率的变化有关, 即曲线斜率变化大处兴趣点多。

假设共有 c 个兴趣点, 令SHGC图像中横截面的第 i 对兴趣点对之间的像素差为 n_i , 横截面 π 中相同 x 轴位置对应第 i 对兴趣点对之间的像素差为 n_2 , 因为第2)步已经考虑了两端点情况, 不再考虑两端切线情况。目标函数为:

$$\min \sum_{i=1}^{(c-2)/2} (n_i - n_2)^2 \quad (11)$$

通过该目标函数优化 λ 的值, 重复第2)步与第4)步, 如果式(11)的误差平方和达到最小时, 则停止上述循环; 如果最终提取的精度还是不够, 可以再微调第3)步的平移分量, 从而一定程度上削减射影变换带来的细微误差。寻找满足式(11)的解, 就是寻找一组 (λ, θ) , 使式(11)达到最小值。而 θ 角可以通过上述方法计算一个初值, 可以看成是微调, 而对于 λ 就可以看成是粗调。

有两种方法可以确定 (λ, θ) 的值: 1) 因为 λ 与 θ 无关, 且 θ 已有相对正确的初值, 因此, 先将 θ 的初值代入算法中, 通过改变 λ 使式(11)达到最小化, 之后再微调 θ 值。2) 结合神经网络与遗传算法寻找

最优解, λ 与 θ 是神经网络的两个输入。本文实验采用的是第1)种方法。

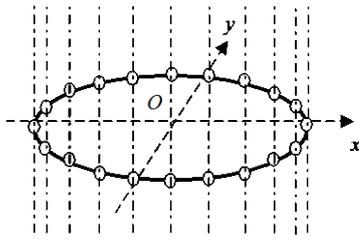


图2 SHGC横截面上选择兴趣点示意图

5) 将第4)步中得到的横截面仿照文献[2]中的方法按照一定比例进行缩放, 并沿对称轴(z轴)方向进行逐个兴趣点的迭代, 缩放比例可以通过第1)步中得到的SHGC子午线轮廓结构得到, 从而得到整个SHGC的三维信息。

3) 提取SHGC物体的纹理。如果要提取SHGC物体的表面的所有纹理, 一幅图像显然是不够的, 需要多幅图像才能全部覆盖SHGC的表面。提取纹理后将图片作处理, 使之更加清晰。本文实验在提取纹理时, 如果是不清晰或者看不到的纹理, 则用最近的纹理网格中的纹理代替, 提取的纹理就存在一定程度上的失真, 又因为实验摄像机是比较普通的摄像头, 所以贴上纹理的效果不是很好。

3 实验

因为本文的算法对摄像机并没有性能等方面的要求, 本文采用的实验摄像机是普通的摄像头。图3是一个瓷器大茶壶的实验原图, 而图4是从图3中截取的目标物体。



图3 瓷器大茶壶



图4 瓷器大茶壶(原图与目标物体提取图)

将原大茶壶沿其子午线提取出轮廓, 如图5所示。实际实验中选择了60组数据。



图5 大茶壶子午线提取示意图

图6所示的实验是本文的重点。从大茶壶的横截面提取横截面的三维结构图, 将该横截面与构造的横截面上的兴趣点数据分别以组为单位依次进行匹配。如图6a所示, 在y轴的负方向有两排兴趣点, 是因为靠近摄像机一端的壶壁在图像中成像更宽, 因此, 需要同时考虑大茶壶的内外壁, 再计算它们的加权平均值作为最终结果。最后, 将选取的20组数据在三维空间中校正, 恢复仿射变换的缩放比例因子, 从而恢复原横截面的三维结构。

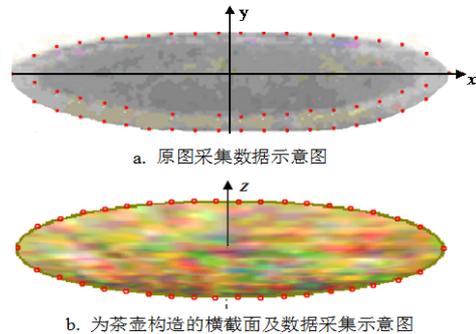


图6 茶壶横截面数据采集与横截面构造的结果示意图

表1 重建横截面图时从源图中采集到的y轴截距数据

左上角	右上角	右下角	左下角
0.0	18.6	-0.0	-19.6
8.4	18.2	-8.4	-19.4
12.4	17.8	-11.6	-18.8
14.0	17.2	-15.2	-18.4
15.6	16.6	-16.8	-17.0
16.4	15.8	-17.2	-16.6
17.0	14.6	-18.4	-15.2
17.8	12.6	-19.0	-11.4
18.2	8.8	-19.4	-8.2
18.6	0.0	-19.6	-0.0

表2 重建横截面图时构造二维横截面得到的y轴截距数据

左上角	右上角	右下角	左下角
0.0	18.8	-0.0	-18.8
8.4	18.6	-8.4	-18.6
11.5	18.1	-11.5	-18.1
13.7	17.5	-13.7	-17.5
15.3	16.6	-15.3	-16.6
16.6	15.3	-16.6	-15.3
17.5	13.7	-17.5	-13.7
18.1	11.5	-18.1	-11.5
18.6	8.4	-18.6	-8.4
18.8	0.0	-18.8	-0.0

本文实验构造的横截面在三维空间中旋转的角度为 78.0° ，即 $\theta=12.0^\circ$ ，且 $\lambda=1.0$ 时达到最优解，其中平移分量为 $\mathbf{0}$ 矩阵，旋转矩阵为单位阵 \mathbf{I} 。表1与表2中的数据分别是达到最优解时原目标物体图中采集的20组数据与构造的横截面中采集的20组数据对。因为计算误差平方和时只需考虑一个坐标轴上的数据，本文实验只考虑y轴方向的数据。另外，因为原图中目标物体偏小，为了方便计算，本文实验先将原图(即所有实验数据)放大5倍后，提取目标物体的横截面三维信息，而表中的数据是将实验结果的数据又缩小到五分之一，与原图数据一致，因此表1与表2中的数据就是像素值。

本文实验按照式(11)的误差平方和为18.620个像素，而平均误差平方和达到0.467个像素。

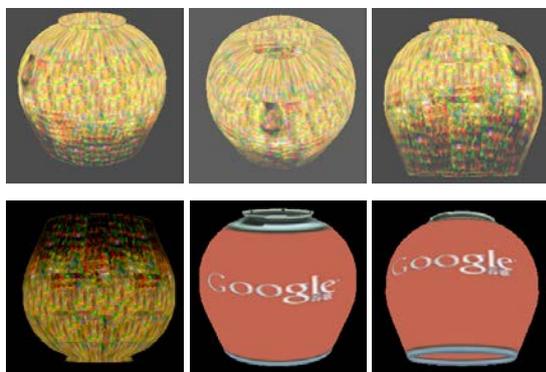


图7 提取大茶壶三维结构后将贴上纹理的效果图

实验的最终结果如图7所示。通过本文设计的算法，将横截面与沿子午线的轮廓结合起来，得到整个大茶壶的三维结构。本文实验在该三维结构上贴了两种纹理，前4幅的纹理是从原图(单幅图像)中提取的纹理，后两幅图的纹理是获取大茶壶的三维结构后，从另一幅表面贴有“Google”字样的杯子上获取的纹理。

4 结论

通过实验中的数据可以看出，实验中平均误差平方和可达到0.467个像素，但是本文的重点并不是仅仅通过该数据来体现的，该数据只是印证了本文的方法的可行性与正确性。因为该数据是原图中采集的数据与构造的横截面上采集的数据进行计算得到的。

文献[8]首次提出了定性视觉的概念^[8]，而本文提出的基于单幅SHGC图像和三维校正的三维结构提取方法也是一种定性视觉信息的提取方法。实际上，根本无法从量化的角度给出评判，因为从单幅图像中无法定量地获取物体真实的三维信息，只能用于定性分析。

在三维空间中逆向校正的方法，可以很方便并准确地提取SHGC物体横截面的三维信息，因为图像本身就是从三维空间中获取的。本文的创新之处在于，提出了一种新的提取单幅SHGC三维结构信息的方法，尤其是提出了对SHGC物体横截面重建的方法，通过原图中采集的数据与构造的横截面上采集的数据所计算得到的平均误差平方和可以看出，本文方法的可行性与正确性。

参 考 文 献

- [1] PONCE J, CHELBERG D, WALLACE B M. Invariant properties of straight homogeneous generalized cylinders and their contours[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1989, 11(9): 951-966.
- [2] ULUPINAR F, NEVATIA R. Shape from contour: Straight homogeneous generalized cylinders and constant cross section generalized cylinders[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1995, 17(2): 120-135.
- [3] LAVEST J, GLACHET R, DHOME M, et al. Modeling solids of revolution by monocular vision[C]//Proc Int Conf on Computer Vision and Pattern Recognition. Lahaina, Hawaii: [s.n.], 1991: 690-691.
- [4] SATO H, BINFORD T O. Finding and recovering SHGC objects in an edge image[J]. Computer Vision, Graphics, and Image Processing: Image Understanding, 1993, 57(3): 346-358.
- [5] ZISSERMAN A, MUNDY J, FORSYTH D, et al. Class-based grouping in perspective images[C]//Proc Int'l Conf Computer Vision. MIT, USA: [s.n.], 1995: 183-188.
- [6] ZERROUG M, NEVATIA R. Part-based 3D descriptions of complex objects from a single image[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1999, 21(9): 835-848.
- [7] COLOMBO C, ALBERTO D B, PERNICI F. Metric 3D reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2005, 27(1): 99-114.
- [8] ALOIMONOS Y. Purposive and qualitative active vision[C]// Proc. IEEE 10th ICPR Atlantic Citys, USA:[s.n.]. 1990: 346-360.
- [9] DOIGNON C, MICHEL D M. A degenerate conic-based method for a direct fitting and 3-D pose of cylinders with a single perspective view[C]//Proc Int'l Conf on Robotics and Automation. Roma, Italy: [s.n.], 2007: 4220-4225.
- [10] HARTLEY R, ZISSERMAN A. Multiple view geometry in computer vision[M]. Cambridge: Cambridge University press, 2004.
- [11] TSAI R Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV camera and lenses[J]. IEEE Journal of Robotics and Automation, 1987, 3(4): 323-344