基于 SVDD 的三维目标多视点视图建模

丁 昊 李旭东 赵慧洁

(北京航空航天大学 精密光机电一体化技术教育部重点实验室, 北京 100191)

摘 要: 同一目标在不同观察视点下成像后外形可能有较大差异,因此三维目标多视点视图建模是目标识别的关键. 针对该问题,提出了基于支持向量数据描述(SVDD, Support Vector Data Description)方法对目标特征进行描述. 在视点球面上均匀采样获取目标全姿态图像,以 SVDD 方法求取在高维空间内包含尽可能多目标特征向量的最小超球体相关参数,得到数量较少的支持向量将作为目标多视点视图的最佳模型. 对多类目标不同姿态的图像(每类2 592帧),以规则化不变矩描述目标外形特征,进行了建模实验,并通过识别实验验证了所提方法的有效性和可行性.

关 键 词: 多视点建模; 支持向量数据描述; 三维目标识别; 规则化不变矩

中图分类号: TP 391

文献标识码: A **文 章 编 号**: 1001-5965(2012)11-1517-05

Method of multi-view modeling for 3D target based on SVDD

Ding Hao Li Xudong Zhao Huijie

(Precision Opto-mechatronics Technology Key Laboratory of Education Ministry,
Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: Popular 3D target recognition approaches based on image continue to struggle with challenge—viewpoint sensitivity. Multi-view modeling technique for 3D target offers promise for this challenge. A method of modeling 3D target based on support vector data description (SVDD) that could obtain a tight description covering most of the target feature data was proposed. Target's images were captured with uniform grid on the viewing sphere and characterized by feature vectors. The support vectors representing characteristic views were obtained by applying SVDD to optimize the parameters of the minimal hyper-sphere which covers as many feature vectors as possible. The experiments were conducted by applying the proposed method to an image set (each target includes 2592 images) characterized by normalized moment invariants. The results show that the proposed method is effective and feasibility.

Key words: multi-view modeling; support vector data description (SVDD); 3D target recognition; normalized moment invariants

三维目标识别是计算机视觉领域内一个重要的研究方向. 目前在实际应用中获取目标的三维信息往往十分困难,识别三维目标仍然主要通过识别目标二维投影形成的图像完成. 成像(投影)过程导致了部分目标信息丢失,且在不同视点下,复杂目标的外形差异明显,使得快速稳定的三维目标识别难度很大.

通过对三维目标进行建模以形成对目标外形的全面描述,是解决该问题的手段之一. 这就需要研究三维目标的多视点视图建模方法. 直观的想法是将多个视点下对目标图像的描述综合起来作为目标的描述. 然而对于同一目标,其姿态数目无

收稿日期: 2011-07-18; **网络出版时间**: 2012-11-09 17:21
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20121109.1721.017.html
基金项目: 国家自然科学基金资助项目(60802044)
作者简介: 丁 昊(1986-),男,云南昆明人,博士生,17dhao@aspe.buaa.edu.cn.

法穷举,不同描述间也包含了很多冗余信息. 因此,需要在约简冗余描述的同时尽可能保留重要的描述.

本文在综述三维目标多视点视图建模方法的基础上,提出了基于SVDD[1](Support Vector Data Description)的三维目标多视点视图建模方法. 其基本思想是:利用包含一组数据的最小超球体边界来描述数据的分布. 由于表征超球体边界的参数较少,因此可以达到约简数据的目的. 通过对视点球面的均匀采样可以得到三维目标的全姿态图像集,使用SVDD表征从图像中提取的特征向量在特征空间中的分布,得到较少且具有充分代表性的特征向量,作为目标的特征描述.

1 三维目标多视点视图建模方法

在三维欧氏空间中定义坐标系,目标位于坐标原点. 不同视点下目标的成像过程可看作目标以原点为起点向不同方向投影的过程. 在所有投影方向上,单位长度投影方向向量形成一个单位球面,即为视点球面[2],其上每一点对应一个投影方向向量. 沿投影方向 P 的正交投影对应于从视点球面上 P 点看目标.

文献中提及的三维目标多视点视图建模方法可分为两类:

第1类方法是基于目标几何结构和可见关系的视点球面分割算法,主要基于目标图像的几何拓扑特征,最终得到的建模结果是关于目标的多姿态特征视图集. 该类方法主要以文献[3]提出的方向图概念为基础,依据物体几何特征和视点间的变化关系实现目标建模. 如文献[4]提出将划分视点球面的过程转化为求取球面上偶然视点的轨迹;文献[5]提出将求取特征视图集的过程等效为划分方向图空间的过程,依据图像结构图是否同构,用直接分划法和分层分划法求取目标特征视图集;文献[6]提出采用分割合并法分划视点球面;文献[7]提出基于凹凸性和交叉关系的修剪算法改进EEE事件(EEE event)的计算,有效地提高了计算效率. 该类方法构建目标特征视图集时对目标几何结构和特征依赖较大,且在实际应用中较难实现.

第2类方法是基于目标特征相似度的聚类算法,将在视点球面上均匀采样得到的目标图像用特征向量描述,并对特征向量集进行聚类,将最终得到的少量目标特征向量作为建模结果. 如文献[2]采用的基于Zernike矩的 K 均值算法;文献[8]提出的基于形状相似性测度(shock graph matching cost)的聚类算法;文献[9]提出的基于形状上下文(shape context)特征的 K-中心算法;文献[10]采用的基于流形学习的聚类算法. 该类方法因其通用性强,已在工程上得到了较多应用;其缺点在于:①需要知道聚类的类别数,一般难以事先确定;②得到的特征视图在进行最近邻分类时无阈值,只能强制划分到某一类中,不能拒绝数据.

基于SVDD的三维目标多视点视图建模方法通过求解在特征空间包含所有特征点的柔性超球体边缘来描述目标性质,可以较好地克服上述两类算法的缺点.

2 基于SVDD的建模方法

2.1 基本原理

对视点球面均匀采样,获取三维目标的全姿态图像集. 采用合适的特征描述图像中的目标,可得到关于该目标的特征向量集合 $\{\boldsymbol{x}_i \mid i=1,2,\cdots,N\}$,$N$ 为获取的目标图像数. 定义包含该集合且封闭的超球体,该超球体由球心 $\boldsymbol{a}$ 和半径 R 表征. SVDD算法通过最小化 R^2 来求取包含尽可能多数据点的最小超球体,以对该特征向量集进行描述. 对于相同目标不同姿态下对应的特征向量而言,它们将包含于超球体的内部.

求取最小超球体的过程可表述为以下优化问题:

$$\min(F(R,\boldsymbol{a}) = R^2) \tag{1}$$

约束条件为

$$\| \boldsymbol{x}_i - \boldsymbol{a} \|^2 \leqslant R^2 \quad \forall i \tag{2}$$

当特征向量集中出现异常点时,数据点到球心的距离可能不严格小于半径 R,因此,引入松弛变量 $\xi_i \geqslant 0$,上述问题变为

$$\min(F(R,\boldsymbol{a}) = R^2 + C\sum_i \xi_i) \tag{3}$$

约束条件为

$$\| \boldsymbol{x}_i - \boldsymbol{a} \|^2 \leqslant R^2 + \xi_i \quad \forall i \tag{4}$$

$$\xi_i \geqslant 0 \quad \forall i \tag{5}$$

其中,参数 C 控制球体体积和误差的折衷,并作为罚项,控制 ξ_i 的大小和超球体边缘的柔性.

使用拉格朗日乘子法,将式(4),式(5)代入式(3),得

$$L(R,\boldsymbol{a},\alpha_i,\gamma_i,\xi_i) = R^2 + C\sum_i \xi_i - \sum_i \alpha_i \{R^2 + \xi_i - (\| \boldsymbol{x}_i \|^2 - 2\boldsymbol{a}\cdot\boldsymbol{x}_i + \| \boldsymbol{a} \|^2)\} - \sum_i \gamma_i \xi_i \tag{6}$$

其中,$\alpha_i \geqslant 0, \gamma_i \geqslant 0$ 为拉格朗日乘子.

为求解式(6)的极小值,令该泛函对 $R, \boldsymbol{a}, \xi_i$ 偏导数为 0,可得

$$\frac{\partial L}{\partial R} = 0 \rightarrow \sum_i \alpha_i = 1 \tag{7}$$

$$\frac{\partial L}{\partial \boldsymbol{a}} = 0 \rightarrow \boldsymbol{a} = \sum_i \alpha_i \boldsymbol{x}_i \tag{8}$$

$$\frac{\partial L}{\partial \xi_i} = 0 \rightarrow C - \alpha_i - \gamma_i = 0 \tag{9}$$

由式(7),式(8)可见,超球体球心为所有特征向量的凸组合. 由于求解式(6)相对于 $R, \boldsymbol{a}, \xi_i$ 的最小化问题可以转化为求解其对偶问题的最大化问题,将式(7) ~ 式(9)代入式(6)可有

$$L = \sum_i \alpha_i (\boldsymbol{x}_i \cdot \boldsymbol{x}_i) - \sum_{i,j} \alpha_i \alpha_j (\boldsymbol{x}_i \cdot \boldsymbol{x}_j) \tag{10}$$

约束条件为

$$\sum_i \alpha_i = 1 \quad 0 \leqslant \alpha_i \leqslant C \tag{11}$$

根据 KKT(Karush-Kuhn-Tucker)条件有

$$\alpha_i (R^2 + \xi_i - \| \boldsymbol{x}_i - \boldsymbol{a} \|^2) = 0 \tag{12}$$

$$\gamma_i \xi_i = 0 \tag{13}$$

当 $\alpha_i > 0$ 时,有 $R^2 + \xi_i - \| \boldsymbol{x}_i - \boldsymbol{a} \|^2 = 0$,可由式(9)推出 $\| \boldsymbol{x}_i - \boldsymbol{a} \|^2 \geqslant R^2$,对应的特征向量将位于超球体边缘附近,成为数据描述的必要成分. 这些特征向量称为数据描述的支持向量. 最小超球体的半径由 $0 < \alpha_i < C$ 对应的特征向量与球心的距离确定.

对于一个新特征向量 $\boldsymbol{z}$ 而言,如果它到球心的距离小于或等于超球体半径,则这个对象被接受,否则被拒绝. 即

$$\| \boldsymbol{z} - \boldsymbol{a} \|^2 = (\boldsymbol{z} \cdot \boldsymbol{z}) - 2 \sum_i \alpha_i (\boldsymbol{z} \cdot \boldsymbol{x}_i) + \sum_{i,j} \alpha_i \alpha_j (\boldsymbol{x}_i \cdot \boldsymbol{x}_j) \leqslant R^2 \tag{14}$$

上述方法仅能适用于接近于球形分布的数据集合,在其他情况下这种超球体得到的描述不够紧实. 由于在式(10)和式(14)中,数据点之间的关系以内积形式表达,可用核函数 $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ 代替内积以得到更加灵活的柔性数据描述. 则式(10)变为

$$L = \sum_i \alpha_i K(\boldsymbol{x}_i \cdot \boldsymbol{x}_i) - \sum_{i,j} \alpha_i \alpha_j K(\boldsymbol{x}_i \cdot \boldsymbol{x}_j) \tag{15}$$

判断新特征向量 $\boldsymbol{z}$ 是否被接受时,判断其是否满足:

$$K(\boldsymbol{z} \cdot \boldsymbol{z}) - 2 \sum_i \alpha_i K(\boldsymbol{z} \cdot \boldsymbol{x}_i) + \sum_{i,j} \alpha_i \alpha_j K(\boldsymbol{x}_i \cdot \boldsymbol{x}_j) \leqslant R^2 \tag{16}$$

若满足则接受,反之则拒绝.

当特征空间中分布有多个目标的特征向量时,若单独为某一目标的特征向量建立超球体,不同目标的超球体间可能会发生交叠. 在这种情况下,可将某一类目标的特征向量当作正例,将其他类目标的特征向量当作反例,用正反例联合优化前述问题,以获得最优的超球体.

用 i, j 表示正例下标,用 l, m 表示反例下标,则式(3)所述优化问题变为

$$\min(F(R, \boldsymbol{a}, \xi_i, \xi_l) = R^2 + C_1 \sum_i \xi_i + C_2 \sum_l \xi_l) \tag{17}$$

约束条件为

$$\| \boldsymbol{x}_i - \boldsymbol{a} \|^2 \leqslant R^2 + \xi_i \quad \forall i \tag{18}$$

$$\| \boldsymbol{x}_l - \boldsymbol{a} \|^2 \geqslant R^2 - \xi_l \quad \forall l \tag{19}$$

$$\xi_i \geqslant 0 \quad \xi_l \geqslant 0 \quad \forall i, l \tag{20}$$

同样,使用拉格朗日乘子法处理,最终得

$$L = \sum_i \alpha_i (\boldsymbol{x}_i \cdot \boldsymbol{x}_i) - \sum_l \alpha_l (\boldsymbol{x}_l \cdot \boldsymbol{x}_l) - \sum_{i,j} \alpha_i \alpha_j (\boldsymbol{x}_i \cdot \boldsymbol{x}_j) - \sum_{l,m} \alpha_l \alpha_m (\boldsymbol{x}_l \cdot \boldsymbol{x}_m) + 2 \sum_{l,j} \alpha_l \alpha_j (\boldsymbol{x}_l \cdot \boldsymbol{x}_j) \tag{21}$$

约束条件为

$$\sum_i \alpha_i - \sum_l \alpha_l = 1 \quad 0 \leqslant \alpha_i, \alpha_l \leqslant C \tag{22}$$

经优化求解,即可得到最优超球体的相关参数.

由于 SVDD 和 SVM(Support Vector Machine)在数学上具有相似的表达形式,SVDD 问题的优化求解可以采用和 SVM 的求解一样的方法,如较先进的 CVM(Core Vector Machine)算法[11],Cutting-plane 算法[12],Condensed SVM 算法[13]等.

2.2　目标特征描述

本文以综合仿射不变矩和几何不变矩形成的七维不变矩描述目标外形特征[14],并在特征空间中使用 SVDD 对其进行描述.

文献[15]基于代数不变性理论提出了仿射不变矩,常用的仿射不变矩可由式(23) ~ 式(25)求出:

$$I_1 = \frac{(\mu_{20}\mu_{02} - \mu_{11}^2)}{\mu_{00}^4} \tag{23}$$

$$I_2 = \frac{\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{03}\mu_{21}\mu_{12}}{\mu_{00}^{10}} + \frac{4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{12}^2\mu_{21}^2}{\mu_{00}^{10}} \tag{24}$$

$$I_3 = \frac{\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12})}{\mu_{00}^7} + \frac{\mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)}{\mu_{00}^7} \tag{25}$$

其中,μ_{pq}表示图像的 $p+q$ 阶中心矩.

文献[16]基于图像的一系列规格化中心矩提出了几何不变矩,式(26)是用于与上述仿射不变矩进行综合的几何不变矩前 4 个分量:

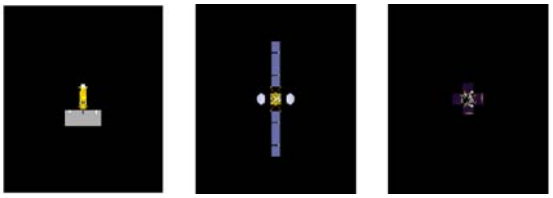
$$\left.\begin{aligned}\Phi_1 &= \eta_{20} + \eta_{02}\\ \Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2\\ \Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2\\ \Phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2\end{aligned}\right\} \quad (26)$$

其中,η_{mn}表示图像的 $m+n$ 阶规则化中心矩.

为解决上述不变矩间数量级差异,需要对上述不变矩进行规则化[17]运算并最终得到七维规则化不变矩.

3 建模实验

首先,获取测试目标的全姿态图像. 将目标置于视点球中心,在视点球面上均匀采样,每个目标将得到2592 帧不同姿态下的图像. 图 1 为目标实例图.



a 目标 1　　b 目标 2　　c 目标 3

图 1 目标图像实例

以图 1 中目标 1 为例,用前述目标特征描述方法计算其全姿态图像的特征向量,得到 2 592 个七维特征向量. 由于七维数据无法可视化显示,采用 PCA(Principle Component Analysis)方法将其进行二维投影,图 2 为其特征在二维空间中的分布图.

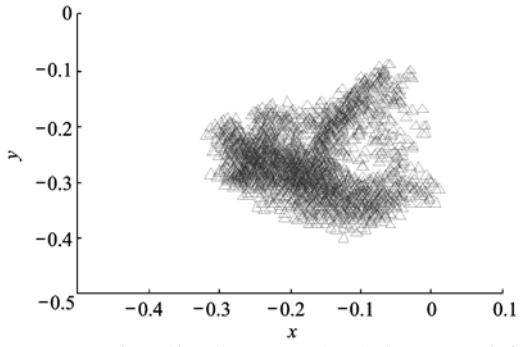


图 2 目标 1 特征向量于二维空间投影后的分布图

采用本文所提基于 SVDD 的三维目标多视点视图建模方法描述上述特征向量集,C 取为 10(此时所得结果紧密性较好),优化求解后,共得到 11 个支持向量,其在总的特征向量集中的分布见图 3. 可以看到,支持向量很好地描述了原特征向量在特征空间中的分布.

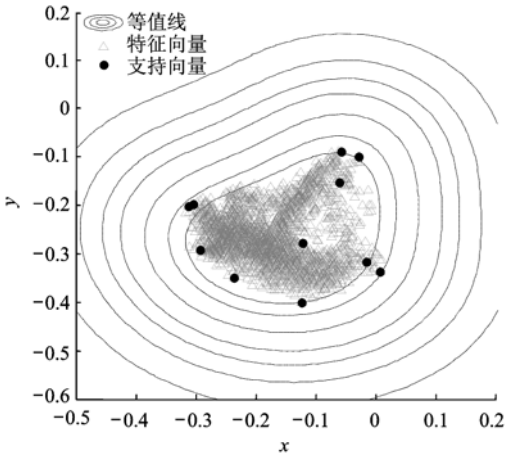


图 3 基于 SVDD 方法对目标 1 建模结果图

采用本文所提基于 SVDD 的三维目标多视点视图建模方法对 3 类目标建模,并进行识别实验. 建模时,采用带有正反例标签的 SVDD 建模方法,训练出其中两类目标的最优超球体参数,并采用级联方式连接两个分类器.

测试图像为目标姿态随机变化时的图像,每类 400 帧. 3 类目标的识别率及对应支持向量数如表 1 所示. 可以看出,基于 SVDD 的三维目标多视点视图建模具有较好的识别效果.

表 1 3 类目标识别率统计表

目 标	测试帧	识别率/%	支持向量数
1	400	87.5	26
2	400	85.8	63
3	400	71.5	

实验中采用分类器级联方式,识别率为两个分类器识别率的乘积,这将影响识别结果. 可尝试采用其他分类方法,如文献[18]提出的多类 Extended SVDD、文献[19]提出的多类 SVDD 模型等.

4 结 论

本文针对三维目标多视点视图建模问题,提出了基于 SVDD 的三维目标多视点视图建模方法,并以规则化不变矩描述目标外形特征,使用该方法对三维目标多姿态图像集建模. 该方法得到的目标描述数量小,冗余信息少,又可以较好地描述目标特征向量在特征空间中的分布. 识别实验结果表明本文方法的有效性和可行性.

参考文献(References)

[1] Tax D, Duin R. Support vector data description [J]. Machine Learning, 2004, 54(1): 45-66

[2] 彭辉,黄士科,陶琳,等. 基于 Zernike 矩的三维目标多视点特性视图建模[J]. 红外与激光工程,2005,34(3):292-296
Peng Hui, Huang Shike, Tao Lin, et al. Multi-view modeling of 3D target based on Zernike moments [J]. Infrared and Laser Engineering, 2005, 34(3): 292-296 (in Chinese)

[3] Koenderink J, Doom A. The internal representation of solid shape with respect to vision [J]. Biological Cybernetics, 1979, 32(4): 211-216

[4] Gigus Z, Malik J. Computing the aspect graph for line drawings of polyhedral objects [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(2): 113-122

[5] Bowyer K, Sallam M, Eggert D, et al. Computing the generalized aspect graph for objects with moving parts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(6): 605-610

[6] Roy S, Chaudhury S, Banerjee S. Aspect graph construction with noisy feature detectors [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2003, 33(2): 340-351

[7] Yu X, Ma H, You S, et al. A solution to efficient viewpoint space partition in 3D object recognition [C]// International Conference on Image and Graphics. Xi'an, Shaanxi, China: IEEE Computer Society, 2009: 565-570

[8] Cyr C, Kimia B. 3D object recognition using shape similarity-based aspect graph [C]// International Conference on Computer Vision. New York, USA: IEEE Computer Society, 2001: 254-261

[9] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522

[10] Pang B, Ma H. An effective way of 3D model representation in recognition system[C]// International Conference on Multimedia and Signal Processing. Guilin, Guangxi, China: IEEE Computer Society, 2011: 107-111

[11] Tsang I W, Kwok J T, Cheung P M. Core vector machines: fast SVM training on very large data sets [J]. Journal of Machine Learning Research, 2005, 6: 363-392

[12] Joachims T, Finley T, Yu C J. Cutting-plane training of structural SVMs [J]. Machine Learning, 2009, 77(1): 27-59

[13] Dung D N, Matsumoto K, Takishima Y, et al. Condensed vector machines: learning fast machine for large data [J]. IEEE Transactions on Neural Network, 2010, 21(12): 1903-1914

[14] 张劲锋,蔡伟. 基于组合不变矩的空间目标识别[C]//全国第十二届空间及运动体控制技术学术会议论文集. 桂林:中国自动化学会空间及运动体控制专业委员会,2006:490-494
Zhang Jinfeng, Cai Wei. Space object recognition based on combined moment invariants [C]//The 12th National Conference on Space and Vehicle Control Technology. Guilin: China Association of Automation, 2006: 490-494 (in Chinese)

[15] Flusser J, Suk T. Pattern recognition by affine moment invariants [J]. Pattern Recognition, 1993, 26(1): 167-174

[16] HU M K. Visual pattern recognition by moment invariants [J]. IRE Transactions on Information Theory, 1962, 8: 179-187

[17] 张天序. 成像自动目标识别[M]. 武汉:湖北科学技术出版社,2005:18-84
Zhang Tianxu. Automated recognition of imaged targets [M]. Wuhan: Hubei Science and Technology Press, 2005: 18-84 (in Chinese)

[18] Mu T, Nandi A K. Multiclass classification based on extended support vector data description [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2009, 39(5): 1206-1216

[19] Yang L, Ma W, Tian B. New multi-class classification method based on the SVDD model [C]// International Symposium on Neural Networks. Guilin, Guangxi, China: Springer, 2011: 103-112

(编　辑:刘登敏)