

基于 K 均值聚类和多示例学习的图像检索方法

温超,耿国华,李展

(西北大学 信息科学与技术学院,西安 710069)

(superheat@163.com)

摘要:针对基于对象的图像检索问题,利用 K 均值(K -means)聚类,提出了一种新的基于多示例学习(MIL)框架的图像检索算法 KP-MIL。该算法在正包和负包组成示例集合聚类,获取潜在正示例代表和包结构特性数据,然后利用径向基核分别度量两者的相似性,最后利用 α 因子均衡两者相似性对核函数结果的影响。在标准对象图像检索集 SIGVAL 上进行实验,实验结果表明,该方法是有效的且性能优于其他同类方法。

关键词:图像检索;多示例学习; K 均值聚类;径向基核; α 因子

中图分类号: TP391.41 **文献标志码:** A

Image retrieval based on K -means clustering and multiple instance learning

WEN Chao, GENG Guo-hua, LI zhan

(School of Information Science and Technology, Northwest University, Xi'an Shaanxi 710069, China)

Abstract: Aiming at the problem of object-based image retrieval, a novel algorithm named KP-MIL was proposed, which worked in the Multiple Instance Learning (MIL) framework. Firstly, this algorithm clustered the instances in positive set and negative set, and found the potential positive instance and bag structure. Then an α coefficient was introduced to trade off between positive instance and bag's similarity. Experiments on SIGVAL dataset show that this algorithm is feasible, and the performance is superior to other MIL algorithms.

Key words: image retrieval; Multiple Instance Learning (MIL); K -means clustering; Radial Basis Function (RBF); α coefficient

0 引言

随着多媒体技术的飞速发展和 Internet 技术的普及,数字图像的种类和数量呈爆炸式增长。为了能够从浩瀚的图像数据库中快速、准确地找到用户所需内容,基于内容的图像检索(Content Based Image Retrieval, CBIR)技术得到了广泛关注^[1]。而基于对象的图像检索(Object Based Image Retrieval, OBIR)只关心那些包含特定对象的图像,由于 OBIR 更适合用户检索需求,从而成为 CBIR 新的研究热点^[2]。

如对图像进行分割,将图像与分割区域分别看做包和示例,包含有检索对象的图像对应的包标为正包,否则标记为负包。利用已标记正负图像学习分类器并进行检索,则基于对象图像检索就变成了多示例学习(Multiple Instance Learning, MIL)问题^[3]。

Zhang 等人^[3], Rahmani 等人^[4]基于多样性密度(Diverse Density, DD)函数,利用 MIL 对 OBIR 进行研究。Chen 等人提出了 DD-SVM^[5]与 MILES^[6]两个经典 MIL 算法。通过构造空间,并将多示例包非线性投影并嵌入到构造空间中,利用支持向量机(Support Vector Machine, SVM)求解 MIL 分类问题。

Dietterich 等人于 1997 年研究药物活性预测问题时,第一次提出了 MIL 问题。在研究中,发现分子存在多种低能形状,而药物实验只能辨别分子参与制药能力,无法确定何种具体形状。针对此问题, Dietterich 将分子看成包,分子的低能形状作为包中的一个示例。在该问题中,如果包标记为负,则

包内所有示例标记都为负;如果包标记为正,则包中必然存在正示例,但是不能确定是哪个或哪些。正示例标记的不确定就形成了 MIL 的难点,也就是歧义性问题。针对 MIL 问题, Dietterich 等人利用轴平行矩形(Axis-Parallel Rectangle, APR)算法解决^[7]。

本文在 MIL 框架下,提出了一种新的对象图像检索方法称之为 KP-MIL (Finding positive instance using K -means clustering to solve MIL)。该方法首先通过在正包和负包示例集合内进行聚类,寻找包内潜在正示例代表和包结构特性数据,以解决 MIL 歧义性问题,然后利用径向基核(Radial Basis Function, RBF)分别度量两者的相似性,最后利用因子 α 平衡两者对包相似性的影响,构造核函数。在基准测试集 SIVAL 上实验,验证了 KP-MIL 算法的有效性。

1 KP-MIL 算法

已有 APR^[7]算法通过贪心式算法找出一个覆盖了每个正包中至少一个示例的最小矩形,然后利用该矩形挑选出最具有区别能力的一组属性,在此基础上对最可能出现正包示例的矩形边界不断进行扩展;而 DD^[3]则利用示例独立同假设和贝叶斯公式将密度函数问题转化为一个函数优化问题,最后在示例空间寻找满足 DD 点,该点其邻域正包最多且负包较远。这两种方法都隐含着要求示例组成空间中正包的潜在正示例具有空间聚集性。基于此认识,本文通过聚类算法获取潜在正示例中心和包内潜在正示例代表,然后构造适合

收稿日期: 2010-12-07; **修回日期:** 2011-02-21。 **基金项目:** 国家自然科学基金资助项目(60873094); 高等学校博士学科点专项科研基金资助项目(200806970014); 陕西省教育厅自然科学基金资助项目(2010JK852)。

作者简介: 温超(1978-),男,陕西三原人,讲师,博士研究生,CCF 会员,主要研究方向:图像检索、模式识别; 耿国华(1955-),女,山东蓬莱人,教授,博士生导师,主要研究方向:模式识别、智能信息处理; 李展(1973-),男,陕西户县人,讲师,主要研究方向:Web 数据挖掘、图像检索。

MIL 问题的核函数,最后用 SVM 进行分类检索。

1.1 获取潜在正示例代表和核函数构建

假定 MIL 问题中,有 s 个训练图像 $B_i (i = 1, \dots, s)$ 。其中有 m 个正包, n 个负包,满足 $s = m + n$ 。对包中示例有 $x_{ij} \in B_i$, $j = 1, \dots, l, y_i \in \{0, 1\}$ 是图像 B_i 的隐藏标记。MIL 就是利用已有训练图像和标记 $B_i = \{(B_i, y_i), \dots, (B_s, y_s)\}$ 学习分类函数,以对新图像 $B_u = \{B_1, \dots, B_l\}$ 进行正确检索。

将正包中的示例排列在一起组成集合 X^+ , 将负包中示例排列在一起组成集合 X^- 。然后在正示例集合 X^+ 和负示例集合 X^- 分别进行 K 均值聚类, 获取正示例聚类中心点 $O^+ = \{O_1^+, O_2^+, \dots, O_k^+\}$ 和负示例聚类中心点 $O^- = \{O_1^-, O_2^-, \dots, O_k^-\}$ 。如果 B_i 包中示例在正聚类的范围内, 则根据正示例的空间聚集性, 那么该示例就有可能是该包的正示例候选点。显然, 哪个正示例聚类中心点范围所包含示例越多, 那么它就越有可能成为真正的正示例。基于此认识, 正示例聚类点集合内包含示例最多的聚类点就是潜在的正示例中心。如果包 B_i 存在示例在正示例中心的范围内, 则该示例就是潜在正示例。在包内如没有示例在正示例中心范围内, 则取和中心距离最近的点作为潜在正示例。在包内有多个示例落在此区域的情况下, 取与潜在的正示例中心距离最近的点作为潜在正示例代表点。

通过聚类形成的向量 $O = O^+ \cup O^-$ 表达了正负向量的数据稠密问题。计算包中示例到聚类向量 O 的距离, 该距离就表征了包中示例和聚类中心的相似度。通过将这些距离进行相似性比较, 就可判断包结构相似性。通过计算包内示例到聚类向量 O 的最小距离 X 和最大距离 Y , 这样就获取了包的结构。

本文选用 K-means 作为聚类方法, 它是最常用的一种聚类算法。K-means 是 MacQueen^[8] 提出的一种划分式聚类算法, 通过迭代更新聚类中心直至收敛。该算法的核心思想是找出 k 个聚类中心 $\{o_1, o_2, \dots, o_k\}$, 使得每一个数据点 X_j 与其最近的聚类中心 o_c 的平方距离和最小化, 其平方和公式为: $J = \sum_{j=1}^k \sum_{i=1}^n \|X_i^{(j)} - o_j\|^2$ 。

下面分析核函数的构建, 假定有两个包 B_i 和 B_j , 其示例为 $B_i = \{x_{i1}, \dots, x_{im}\}$, $B_j = \{x_{j1}, \dots, x_{jm}\}$ 。使用 K-means 聚类分别在 X^+ 和 X^- 聚类, 获取聚类中心为 $O^+ = \{O_1^+, O_2^+, \dots, O_k^+\}$ 和 $O^- = \{O_1^-, O_2^-, \dots, O_k^-\}$ 。假定包含 O_i^+ 包含的正示例数最多, 则 O_i^+ 就为潜在正示例聚类中心。如 $x_{ip} = \operatorname{argmin}_i \|x_u - O_i^+\|$, $x_{jq} = \operatorname{argmin}_j \|x_{js} - O_i^+\|$, 则 x_{ip}, x_{jq} 就分别为 B_i 和 B_j 的潜在正示例代表。新生成的包结构向量为 $y_i^l = \min_{l=1-2k, t=1-m} \|x_u - O_l\|$, $y_j^l = \min_{l=1-2k, t=1-m} \|x_u - O_l\|$, $O = O^+ \cup O^-$ 。将包 B_i, B_j 的潜在正示例代表 x_{ip}, x_{jq} 和其包结构 y_i^l 和 y_j^l , 分别用核函数 k_1 和 k_2 度量相似性, 最后再用 α 平衡两者对包相似性的影响。则最终包相似性结果定义为 $k_{ij} = \alpha k_1(x_{ip}, x_{jq}) + (1 - \alpha) k_2(y_i^l, y_j^l)$, $1 > \alpha > 0$ 。

定理 1 多个核函数的正系数线性组合仍然是核函数。

证明 设 k_i 为核函数, $\alpha_i \geq 0 (i = 1, \dots, n)$, 则需证明 $K = \sum \alpha_i k_i$ 仍然为核函数。如果 k_i 为核函数, 则必然存在矩阵对称半正定矩阵 M_i , 显然 $\sum \alpha_i M_i$ 仍为对称半正定矩阵。

命题得证。

从定理 1 可知, 我们所定义的包相似性函数 k 仍为核函数。以上获取潜在正示例代表和核函数构建过程如算法 1 所

示。

算法 1 核函数构建。

输入: MIL 训练图像和标记 $\{(B_1, y_1), \dots, (B_s, y_s)\}$, MIL 待检索图像 $B_u = \{B_1, \dots, B_l\}$, 聚类数 K , 比例系数 α ;

输出: 核函数矩阵 k ;

初始化: $k = 0$;

- 1) 将标记 $y_i = 1$ 的包 B_i 排列在一起组成集合 X^+ ;
- 2) 在集合 X^+ 中任意选取 K 个示例 $O^+ = \{x_i\}_{i=1}^k$ 作为初始聚类中心;
- 3) 计算每个示例 x 到聚类中心 O^+ 的距离, 把示例 x 分配到距离最近的聚类中;
- 4) 利用 K-means 算法计算 K 个聚类的中心;
- 5) 与前一次计算得到的 K 个聚类中心比较, 如果聚类中心发生变化, 转 3), 否则转 6);
- 6) 计算 O^+ 中每个聚类中心所包含的示例数, 包含示例数最多的 O_i^+ 为潜在正示例中心;
- 7) 令 $B = B_l \cup B_u$, 对集合中每个包 B_i , 计算其到潜在正示例中心 O_i^+ 的距离, 距离最短的 x_{ip} 就为潜在正示例中心;
- 8) 与上面步骤类似, 生成负包的聚类中心 O^- ;
- 9) 令 $O = O^+ \cup O^-$, 对集合 B 中每个包 B_i , 计算其到集合 O 中每个元素 O_j 最短的距离 $y_i^l = \min_{l=1-2k, t=1-m} \|x_u - O_l\|$, 组成向量 Y ;

10) 任取集合 $B = B_l \cup B_u$ 中的两个包 B_i 和 B_j , 计算其核函数 $k_{ij} = \alpha k_1(x_{ip}, x_{jq}) + (1 - \alpha) k_2(y_i^l, y_j^l)$;

11) 输出核矩阵 k 。

1.2 利用 SVM 分类

SVM 基于最大化 margin 思想, 寻找分类超平面的最大间隔。其目标函数如下:

$$\min \varphi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + c \sum \xi_i$$

$$\text{s. t. } y_i(\omega^T \varphi(x_i) + b) \geq 1 - \xi_i \quad (1)$$

其中: $\{x_i, y_i\} (i = 1, \dots, m)$ 为训练样本, m 为训练样本总数; ω 为分类面法向量, ξ_i 为松弛变量; c 为惩罚因子, 如 $c = \infty$ 则为线性可分情况, 否则为线性不可分情况。

通过对偶变形, 其优化问题变为:

$$\max \psi(\alpha) = \sum \alpha_i - \frac{1}{2} \sum \sum \alpha_i \alpha_j y_i y_j k(x_i, x_j)$$

$$\text{s. t. } \sum \alpha_i y_i = 0, 0 \leq \alpha_i \leq c \quad (2)$$

这里 k 为核函数, 如求得优化函数 (2) 的最优解为 α_i^* 和 b^* , 则分类函数为:

$$f(x) = \operatorname{sgn}(\sum \alpha_i^* y_i k(x_i, x) + b^*) \quad (3)$$

最后将本文提出的 KP-MIL 对象图像检索算法步骤总结如下。

算法 2 KP-MIL 算法。

输入: MIL 训练图像和标记 $\{(B_1, y_1), \dots, (B_s, y_s)\}$, MIL 待检索图像 $B_u = \{B_1, \dots, B_l\}$, 聚类数 K , 比例系数 α ;

输出: 待检索图像 $B_u = \{B_1, \dots, B_l\}$ 的标记 $\{y_i\}_{i=1}^l$;

- 1) 调用算法 1, 计算核矩阵 k ;
- 2) 训练 SVM 分类器: ① 输入聚类数 K 值和正比例系数 α 值; ② 求解式 (2) 优化问题;
- 3) 利用式 (3) 求解检索图像 B_u 的类别标记 $\{y_i^*\}_{i=1}^l \in \{-1, +1\}$ 。

2 实验结果与分析

2.1 数据集和实验方法

SIVAL 图像集是一个人为设计的对象图像检索标准测试集。它包含 25 类不同物体, 每类 60 幅图像, 共由 1500 幅图像组成。物体分别放置在 6 种不同位置, 10 个不同场景和光照条件下生成。图像被分割成约 30 个不同区域, 提取每个区域的 30 维的颜色、纹理和近邻的底层视觉特征^[4]。

SVM 代码选用 Libsvm (available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>), 选用一对方法处理多类问题, 在每次实验中, 从某类图像中随机选取 10 幅, 标为正包, 其他 24 类的所有图像中随机选取 10 幅标为负包, 其余图像为待检索包。 k_1 和 k_2 都选择径向基核函数 (Radial Basis Function, RBF)。

2.2 KP-MIL 算法的有效性

KP-MIL 算法中, K -means 聚类的聚类数 K , 以及核函数中的正比例系数值 α 必须预先设置。为了验证 K 值和 α 对检索精度的影响, 首先固定 $\alpha = 0.1$, K 以步长为 1, 从 1 变化到 9, 以分析 K 值对检索性能影响。然后获取最佳的 K 值固定, 让 α 值步长为 0.1, 从 0.1 变化到 0.9。基于 green-teabox 类图像, 按照算法 2 实验。图 1 为固定 $\alpha = 0.1$ 情况下, K 值变化检索结果其平均 AUC (Area Under ROC Curve) 值。图 2 为固定 $K = 5$ 情况下, α 值变化检索结果其平均 AUC 值。

如图 1 所示, K 从 1 变到了 9, AUC 值为 0.73 ~ 0.84。在 $K = 5$ 时精度达到最大值 0.84, 由此可见 KP-MIL 算法的性能受 K 影响较大。其主要原因是: 在对象图像检索中, 对象经常和几个特定其他物体具有特定的共生关系。如果聚类数等于共生数, 则聚类后, 特定对象应以较高频率出现, 而其他共生物体则频率较低。因此聚类后, 包括最多示例聚集中心点必然是潜在正示例中心。如图 2 所示, 固定 $K = 5$, 从 0.1 变到了 0.9, AUC 值在 0.81 ~ 0.86, 当 $\alpha = 0.3$ 时, 达到最大值 0.86。

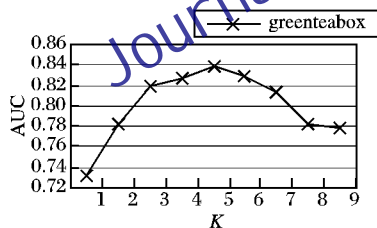


图 1 K 值变化影响 ($\alpha = 0.1$)

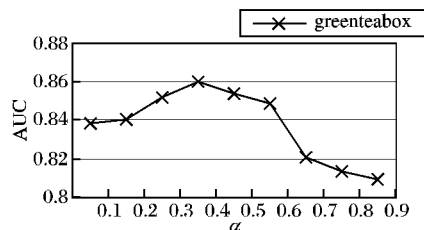


图 2 α 值变化影响 ($k = 5$)

为了进一步验证 KP-MIL 算法的性能 (设置 $K = 5$, $\alpha = 0.3$), 按照 2.1 节所述的实验方法, 将 KP-MIL 和 MILES, ACCIO! 算法进行对比实验, 每个实验重复 30 次, 平均 AUC 值如表 1 所示。表中标号 1 ~ 25, 为 SIVAL 图像集中: FabricSoftenerBox, CheckeredScarf, FeltFlowerRug, CokeCan, WD40Can, AjaxOrange, DirtyRunningShoe, CandleWithHolder, GoldMedal, GreenTeaBox, CardboardBox, SpriteCan,

SmileyFaceDoll, DirtyWorkGloves, StripedNoteBook, DataMiningBook, BlueScrunch, TranslucentBowl, RapBook, Apple, GlazedWoodPot, WoodRollingPin, JuliesPot, Banana, Large-Spoon。

表 1 置信度为 0.95 的 30 次重复实验平均 AUC

标号	KP-MIL	MILES	ACCIO!
1	97.6 ± 0.3	97.1 ± 0.8	86.3 ± 3.3
2	95.2 ± 0.5	93.9 ± 1.3	90.2 ± 1.7
3	95.9 ± 0.9	94.1 ± 0.8	87.9 ± 1.9
4	93.6 ± 0.6	92.8 ± 0.9	81.8 ± 3.1
5	90.9 ± 0.9	88.5 ± 2.1	81.2 ± 2.5
6	91.3 ± 1.8	90.6 ± 2.5	76.8 ± 3.1
7	90.2 ± 1.3	85.6 ± 1.6	83.3 ± 1.8
8	85.6 ± 1.8	84.2 ± 2.2	69.8 ± 2.7
9	82.2 ± 1.4	80.9 ± 2.6	77.9 ± 2.4
10	84.9 ± 2.6	91.3 ± 2.3	86.3 ± 3.5
11	82.8 ± 1.2	82.2 ± 2.8	69.9 ± 2.7
12	82.6 ± 1.1	81.4 ± 2.1	71.1 ± 2.9
13	80.6 ± 1.9	77.8 ± 2.5	77.6 ± 3.8
14	82.3 ± 1.3	78.1 ± 3.5	65.4 ± 1.7
15	76.2 ± 2.1	73.1 ± 3.1	74.8 ± 3.7
16	72.8 ± 2.3	67.7 ± 2.5	71.2 ± 3.1
17	74.6 ± 2.1	73.6 ± 2.2	69.2 ± 3.7
18	73.8 ± 2.9	72.2 ± 3.6	77.8 ± 2.3
19	71.2 ± 2.1	69.1 ± 3.2	62.9 ± 3.2
20	70.1 ± 1.8	61.9 ± 2.5	63.8 ± 1.8
21	70.3 ± 2.2	68.3 ± 3.2	72.8 ± 2.1
22	70.2 ± 2.6	64.8 ± 2.4	62.4 ± 3.7
23	70.0 ± 3.5	78.8 ± 2.7	79.6 ± 2.7
24	67.9 ± 1.8	62.9 ± 2.9	66.1 ± 1.9
25	62.1 ± 1.6	58.9 ± 1.9	57.2 ± 2.7
平均	80.6	78.7	74.5

从表 1 可以看出 KP-MIL 算法总体来说优于其他 MIL 算法, 主要原因是: 本文设计的 K -means 聚类方法, 可以很好的获取潜在正示例中心, 辨别包内潜在正示例代表和其他示例, 利用 RBF 核分别度量相似性, 可以平衡包内各示例对核相似性影响, 从而提高较高的检索功能。而 MILES 则采取将每个包投影到示例构成的空间, 利用新生成投影向量比较相似性, 其性能很大程度受分类图像中包中示例空间分布的影响, 因此在示例空间分布不理想情况下, 其性能表现不佳。ACCIO! 是一种基于 DD 的算法, 存在优化解陷入局部极值问题, 因此最终分类结果也不精确。

3 结语

在基于对象的图像检索问题中, 本文的主要创新点在于, 将 K -means 聚类与 MIL 问题相结合, 提出了一种新的 MIL 算法——KP-MIL 算法, 该算法利用 K -means 聚类获取潜在正示例代表和包结构特性数据, 分别利用 RBF 核度量两者相似性, 并利用 α 因子均衡两者的对核函数结果影响。相对于传统的 DD 算法来说, KP-MIL 算法不但具有更强的全局寻优能力, 而且效率也更高。在 SIVAL 数据集上的实验, 表明该算法是有效的。

由于该算法的性能依赖于聚类参数 K 和比例系数 α , 因此如何快速寻找适合的 K 和 α , 是一个值得进一步研究的问题。
(下转第 1568 页)

通过式(13),解一个 7×7 的线性方程就可以求出摄像机参数 Φ ,再把 Φ 代入式(13)就变换为一个关于 \mathbf{X} 的 3×3 线性方程。整个计算过程中最消耗时间的是左边的表达式 $\mathbf{BC}^{-1}\mathbf{B}^T$ 。表达式中每个元素为:

$$(\mathbf{BC}^{-1}\mathbf{B}^T)_{ij} = \sum_k B_{ik} C_{kk} B_{kj} \quad (17)$$

其中: $\mathbf{B}_{ij} = \frac{\partial r_{ij}}{\partial \theta_i} \frac{\partial r_{ij}}{\partial x_j}$ 是一个 7×3 的矩阵, $\mathbf{C}_{kk} = \sum_i \frac{\partial r_{ik}}{\partial x_k} \frac{\partial r_{ik}}{\partial x_k}$ 是一个 3×3 的矩阵。

4 实验结果

本文算法采用 OpenCV 和 OpenGL 实现。图 1 为输入的分辨率为 640×480 ,多角度获得的 5 幅序列图像。图 2 为重建后的三维点网格模型和纹理映射后三维模型。模型可以在 OpenGL 中用不同的视角旋转。计算出三维坐标重投影误差为:

$$\varepsilon = \sum_{j=1}^n d(\theta_i X_j, x_{ij})^2 \quad (18)$$

其中: d 表示欧氏距离, n 表示特征点的个数。

利用图 1 的序列图像进行了三维重建实验验证,实验结果如表 1 所示。从表 1 可以看到,迭代 20 次,采用鲁棒优化,重投影误差(ε)减少了 61.9%,二维均值误差(e)减少了 40.1%。迭代 40 次,重投影误差减少了 77.3%,二维均值误差减少了 57.0%。迭代 60 次,重投影误差减少了 98.6%,二维均值误差减少了 70.8%。可以看出随着迭代次数的增加,重建误差越来越低,三维重建的鲁棒性得到了显著的提高。

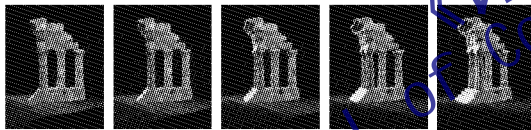


图 1 五幅不同视角的序列图像

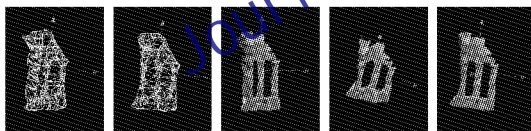


图 2 重建后的三维网格模型和纹理映射模型

表 1 重投影误差对比分析表

迭代次数	鲁棒优化	ε / mm	e / pixel
20	无	2.584	1.256
	有	0.985	0.752
40	无	1.236	0.856
	有	0.281	0.368
60	无	0.975	0.521
	有	0.014	0.152

5 结语

基于序列图像的三维重建相对于其他重建方法具有成本低、操作简单等优点,是目前计算机视觉领域的一个研究热点,但是计算较为复杂,且鲁棒性不够理想,对于大量序列图像更是如此。许多研究通过对算法进行改进来提高鲁棒性,并取得了一定的效果。

本文以两视图重建为基础研究序列图像的三维重建,提出了基于稀疏捆集调整的鲁棒三维重建方法。实验结果表明,本文方法提高了序列图像三维重建的精度和鲁棒性,并真实地再现了物体的三维模型。

参考文献:

- [1] FAUGERAS O, LUONG Q-T. The geometry of multiple images [M]. Cambridge: The MIT Press, 2001: 45-76.
- [2] HEYDEN A, ASTROM K. Euclidean reconstruction from constant intrinsic parameters [C]// Proceedings of International Conference on Pattern Recognition. Washington, DC: IEEE Computer Society, 1996: 339-343.
- [3] FURUKAWA Y, PONCE J. Accurate, dense, and robust multiview stereopsis[J]. IEEE Transactions on Pattern Analysis and Machine Learning, 2010, 32(8): 1362-1376.
- [4] HARTLEY R, ZISSERMAN A. Multiple view geometry in computer vision[M]. Cambridge: Cambridge University Press, 2000: 20-56.
- [5] NISTER D. An efficient solution to the five-point relative pose problem[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(6): 756-770.
- [6] KONOLIGE K, AGRAWAL M. FrameSLAM: from bundle adjustment to realtime visual mapping[J]. IEEE Transactions on Robotics, 2008, 24(5): 1066-1077.

(上接第 1548 页)

参考文献:

- [1] DATTA R, LI J, WANG J Z. Content-based image retrieval approaches and trends of the new age[C]// Proceedings of the 7th International Workshop on Multimedia Information Retrieval, Inconjunction with ACM International Conference on Multimedia. New York: ACM, 2005: 253-262.
- [2] DATTA R, JOSHI D, LI J, et al. Image retrieval: Ideas, influences, and trends of the new age[J]. ACM Computing Surveys, 2008, 40(2): 1-65.
- [3] ZHANG QI, GOLDMAN S A. Content-based image retrieval using multiple-instance learning[C]// Proceedings of the 19th International Conference on Machine Learning. San Francisco: Morgan Kaufmann, 2002: 682-689.
- [4] RAHMANI R, GOLDMAN S A. Localized content-based image retrieval[C]// Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval. New York: ACM, 2005: 227-236.
- [5] CHEN YIXIN, WANG J Z. Image categorization by learning and reasoning with regions[J]. Journal of Machine Learning Research, 2004, 5(8): 913-939.
- [6] CHEN YIXIN, BI JINBO, WANG J Z. MILES: Multiple-instance learning via embedded instance selection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(12): 1931-1947.
- [7] DIETTERICH T G, LATHROP R H, LOZANO-PÉREZ T. Solving the multiple instance problem with axis-parallel rectangles[J]. Artificial Intelligence, 1997, 89(12): 31-71.
- [8] MacQUEEN J. Some methods for classification and analysis of multivariate observations[C]// Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. Los Angeles: University of California Press, 1967: 281-297.