

直推式多视图协同分割

朱云峰* 章毓晋

(清华大学信息科学与技术国家实验室 北京 100084)

(清华大学电子工程系 北京 100084)

摘 要: 多视角拍摄条件下获取到含有同一刚性或静态目标多幅图像后,快速准确地所有图像中分割出该目标是一个尚未被关注的问题。该文首次给出这个问题的数学描述,并借助图模型描述方法,推导了它与传统的基于图割分割问题之间的关系。在求解该问题时,该文提出了迭代式协同直推优化算法,算法思想是将优化目标函数拆分为图像空间和 3 维空间分割两个子优化问题,利用图割和谱聚类分割的方法迭代求解两个子优化问题,并在迭代过程中设计传播、过滤、投票处理引入两个空间标签一致性约束,算法收敛时同时得到图像空间和 3 维空间的分割结果。最后,使用实际图像序列实验得出算法的平均误判率为 3.4%,比较讨论和改进思路一并给出。

关键词: 图像处理;直推式学习;图割;谱聚类分割;多视角协同分割

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2011)04-0763-06

DOI: 10.3724/SP.J.1146.2010.00839

Transductive Co-segmentation of Multi-view Images

Zhu Yun-feng Zhang Yun-jin

(Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China)

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract: Fast and efficient segmentation of rigid or stable object in multi-view images is still a unsolved problem. In this paper, the problem is formulated and the relationships between it and traditional min-cut based segmentation problems are also deduced with Graph representation. To minimize the energy function, a novel algorithm named Interactive Transductive Co-Segmentation (ITC-Seg) is proposed. In ITC-Seg, the function is divided into two sub problems which are solved with graph cuts and spectral segmentation methods. Moreover, propagation, filter, voting methods are introduced into the iteration between multi-view images segmentation and quasi-sparse 3D points segmentation, they are used to combine the sub problems with a global label consistent constrains. Finally, the experiments in several images show the error rate of ITC-Seg is 3.4%, discussions and future improvements of the method are also given.

Key words: Image processing; Transductive learning; Graph cuts; Spectral segmentation; Multi-view co-segmentation

1 引言

随着本地及网络上共享的图像数量急剧增加,获取到特定目标或者类似目标的大量图像数据变得比较简单。随之而来的一个问题是,如何快速准确地从得到的大量图像中分割出相同或相似的目标。由此引发了目前图像分割领域一个研究热点,即协同分割(co-segmentation)^[1]。协同分割中待分割的目标可以是相同或相似的。然而,当多幅图像中背景也存在相同或相似区域时,没有用户的交互,目标

和背景的定义是有歧义的。另外,如果获取到的大量图像中目标是刚性或静态目标(相同目标),基于运动恢复结构的方法^[2]可以准确地从这些图像中恢复出相机的参数和 3 维点云。虽然这些 3 维点不足以生成目标的 3 维几何模型,但其空间信息对图像中目标分割显然是有帮助的,反过来,图像中的目标轮廓又可以用来进行 3 维目标几何建模^[3,4]。

单个图像分割是协同分割的基础,目前,基于图表达的方法是图像分割的一大分支,如图割方法(graph cuts)^[5]、随机游走方法(random walker)^[6]、Grab cut 方法^[7]。其中 Grab cut 方法可以看作是彩色图像中迭代式图割方法。随机游走分割准则可理解为半监督直推式分割^[8],也可解释为交互式的归一

2010-08-09 收到, 2011-01-14 改回

国家自然科学基金(60872084)资助课题

*通信作者: 朱云峰 zhu-yf06@mails.tsinghua.edu.cn

化割方法^[9-11]。在多个视图分割方面,文献[12]使用一幅图像中用户精确提供的目标边界信息,分别建立前景背景局部颜色模式模型和边界模型,并使用马尔科夫随机场模型对新的图像进行像素标记,虽然该文名称为直推式目标分割(transductive object cutout),但并未利用测试图像的信息,因此更接近归纳式学习方法(inductive learning)^[13]。文献[14,15]使用双目摄像机获取的视频,融合深度信息和纹理信息将目标从背景中分割出来,这些方法都隐式地使用了3维空间信息。文献[16,17]提出了利用3维重建点和2维图像进行联合聚类的方法,但要求3维空间点是准稠密的,该方法未考虑有3维空间点的噪声情况,因此错误标记的3维点将导致多幅图像中目标标记错误。此外,该文重点讨论多类聚类问题(即场景多类分类问题)。

受上述研究的启发,本文首次提出一个尚未被研究的问题:两幅及其以上视图中包含刚性或静态单一目标时,在使用基于运动恢复结构的方法获取到相机姿态信息和3维点云后,基于该输入数据,在用户简单交互下,如何快速准确地在图像空间和3维空间分割出该目标,即获取到目标在图像中的轮廓信息及3维空间中目标的稀疏表面点信息。

分割包含3维空间点云上的分割以及图像空间中像素点的分割,有些目标在3维空间是明显分离的,而有些目标虽然在3维空间是相邻的,但它们之间在图像上又存在明显的边界。将两类分割结合起来是本文的研究重点。本文算法的创新之处是:将该问题数学化为对像素点及空间点进行两类分类标记的问题,利用图结构模型,采用半监督直推式学习方法交替地在图像空间和3维空间进行分割,并在图像与图像之间、图像点与3维点之间传递预测标记以及前景和背景的先验模型,最终给出像素点及空间点的标记信息。

2 问题数学描述

首先给出符号定义, $I = \{I^{(k)}\}$ 为多视角图像集, $C = \{C^{(k)}\}$ 为该图像上所有像素点颜色的集合, $P = \{P^{(k)}\}$ 为图像对应的相机投影矩阵集合,其中, $k = 1, \dots, n$, $n \geq 2$ 为图像个数。 $P = \{P_i\}$ 为3维空间点,其中 $P_i = (x_i, y_i, z_i)$, $i = 1, \dots, m$, m 为3维点的个数。 $U_k = \{u_0^{(k)}, u_1^{(k)}, \dots, u_i^{(k)}\}$ 为3维空间点在图像 k 中的投影点集, $u_i^{(k)}$ 为图像中的像素坐标, $c_i^{(k)}$ 为该像素的颜色值,由投影变换关系可知 $u_i^{(k)} = P^{(k)}(x_i, y_i, z_i, 1)^T$, 此处 $i \leq m$ 。 $V_i = \{v_i^{(k)}\}$ 为3维点 P_i 的可见性集合,例如 $V_6 = \{2, 3, 4\}$, 即 P_6 点在序号为 2, 3, 4 的图像中未被遮挡, 分别对应图像中 3 个投影

点 $\{u_6^{(2)}, u_6^{(3)}, u_6^{(4)}\}$ 。 $Y_k = \{y_j^{(k)}\}$ 为图像 k 中像素的标签集合, $y_j^{(k)}$ 为图像 k 中像素 j 的标签, $Y^{(3D)} = \{y_i^{(3D)}\}$ 为3维点标签集合, $y_i^{(3D)}$ 为3维点 P_i 的标签,其中 P, P, V 由基于运动恢复结构的方法^[2]给出。

本文要解决的问题为:当 I, P, P, V, U 已知情况下,求解满足约束的最优 $Y, Y^{(3D)}$ 。首先采用图的描述对该问题进行数学建模,分别在图像空间对每一个图像(k)和3维空间(3D)定义加权图 $\mathcal{G}^{(k)} = (\mathcal{V}^{(k)}, \mathcal{E}^{(k)})$, $\mathcal{G}^{(3D)} = (\mathcal{V}^{(3D)}, \mathcal{E}^{(3D)})$, 如图1所示。图像序列分割和3维空间点分割问题转化为在图 $\mathcal{G}^{(k)}$ 和 \mathcal{G}^{3D} 最优化如下能量函数:

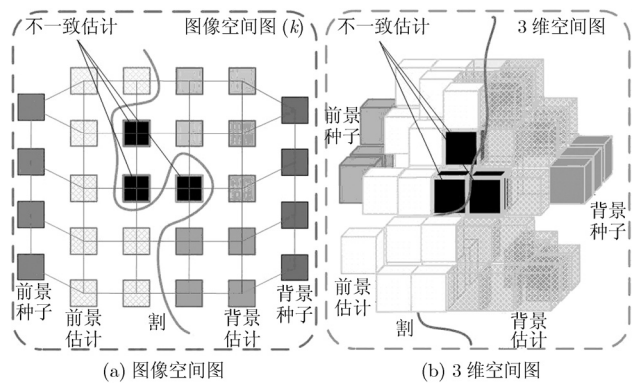


图1 图像空间与3维空间图描述

$$\left. \begin{aligned} & \operatorname{argmin}_{\mathbf{x}=(\mathbf{x}^{(k)}, \mathbf{x}^{(3D)})} E_P(\mathbf{x}^{(k)}, \mathbf{x}^{(3D)}) \\ & = \sum_{(k) \in I} \sum_{e_{ij}^{(k)} \in \mathcal{E}^{(k)}} \omega_{ij}^{(k)} |x_i^{(k)} - x_j^{(k)}|^p \\ & \quad + \sum_{e_{ij}^{(3D)} \in \mathcal{E}^{(3D)}} \omega_{ij}^{(3D)} |x_i^{(3D)} - x_j^{(3D)}|^p, \quad k \in [1, \dots, m] \\ & \text{s.t. } x(\mathcal{O}) = 1, x(\mathcal{B}) = 0, y_i = 1 \text{ if } x_i \geq \frac{1}{2}, 0 \text{ if } x_i < \frac{1}{2} \\ & y_{X_M}^{(3D)} = y_{X_M}^{(k)}, y_i^{(3D)} = y_{u_i}^{(k)} \end{aligned} \right\} \quad (1)$$

式(1)中约束条件为3维点 i 和图像坐标点 u_i 在投影变换时标记 y 应保持一致。针对如何快速有效地逼近这个 NP 问题的全局最优解,下面给出本文提出的迭代式协同分割算法。

3 迭代式协同分割算法

$$\begin{aligned} E_P(\mathbf{x}) \text{ 函数表达中有两个主要部分} \\ E_1 = E_P(\mathbf{x}^{(k)}) & = \sum_{(k) \in I} \sum_{e_{ij}^{(k)} \in \mathcal{E}^{(k)}} \omega_{ij}^{(k)} |x_i^{(k)} - x_j^{(k)}|^p \\ E_2 = E_P(\mathbf{x}^{(3D)}) & = \sum_{e_{ij}^{(3D)} \in \mathcal{E}^{(3D)}} \omega_{ij}^{(3D)} |x_i^{(3D)} - x_j^{(3D)}|^p \end{aligned}$$

这两部分要同时满足的约束是,3维空间点与图像像素点间的投影关系导致的对应标记必须一致。如

图1中黑色块位置,如果仅仅考虑图像空间和3维空间求解子问题,就会出现标记不一致问题,即两个空间中标记预测出现不一致。本文借鉴解决复杂优化目标的对偶分解方法(dual decomposition)^[18]和稀疏编码(sparse coding)^[19]中的交替坐标下降(alternative coordinate descent)方法的思想,首先分别对两部分子函数进行优化求解,然后在迭代过程引入因投影关系产生的标记一致性约束。

E_1 是每一幅图像中标记能量 $E_{1k} = \sum_{e_{ij}^{(k)} \in \mathcal{E}^{(k)}} \omega_{ij}^{(k)} |x_i^{(k)} - x_j^{(k)}|^{p^{(k)}}$ 之和,在不考虑 E_2 时,每幅图像上的 E_{1k} 是独立的。加入由用户提供的种子标记点这个约束,将 E_{1k} 改写为

$$\left. \begin{aligned} E_{1k} &= \sum_{e_{ij}^{(k)} \in \mathcal{E}^{(k)}} \omega_{ij}^{(k)} |x_i^{(k)} - x_j^{(k)}|^{p^{(k)}} \\ &+ \sum_{v_i} \omega_{\mathcal{O}_i}^{(k)} |x_i^{(k)}|^{p^{(k)}} + \sum_{v_i} \omega_{\mathcal{B}_i}^{(k)} |x_i^{(k)} - 1|^{p^{(k)}} \\ \text{s.t. } x^{(k)}(\mathcal{O}) &= 1, x^{(k)}(\mathcal{B}) = 0, y_i^{(k)} = 1 \text{ if } x_i^{(k)} \geq \frac{1}{2}, \\ &0 \text{ if } x_i^{(k)} < \frac{1}{2} \end{aligned} \right\} (2)$$

函数中当 $\forall x_i^{(k)} \in \{0,1\}$, $p^{(k)} = 1$ 时,转化为传统的图割分割^[5]问题。针对彩色图像,本文采用 Grab Cut 算法^[7]对用户任意选择出来的第1幅图像进行交互式分割。当函数中 $\forall x_i^{(k)} \in [0,1]$, $p^{(k)} = 1$ 时,与传统的图割方法^[7]不同之处在于,种子点可以来自于用户交互,也可以是学习算法决策提供,当3维空间点给出分割结果后,3维点上的 $x_p^{(3D)}$ 投影对应图像的像素点 $u_p^{(k)}$ 上的 $x_{u_p}^{(k)}$ 取值可以是 $[0,1]$ 间的任意实数,该值作为像素所属类别的初始概率,在训练前景和背景颜色的高斯混合模型时,就有了初始样本点。选择所有 $x_{u_p}^{(k)} \in [T_o, 1]$ 归入前景集合 $\mathcal{O}_M^{(k)}$, $x_{u_p}^{(k)} \in (0, T_b]$ 归入背景集合 $\mathcal{B}_M^{(k)}$, 所有未标记的点 $x_i^{(k)} = 1/2$, 试验中 T_o 为 0.95, T_b 为 0.05, 以确保置信度很高的点作为训练混合高斯模型的样本点。未与源节点和汇聚节点的连接边权重为 $\omega_{ij}^{(k)} = \exp(-\beta |x_i^{(k)} - x_j^{(k)}|^\alpha \cdot \|c_i^{(k)} - c_j^{(k)}\|^2)$, $\beta > 0, \alpha > 0$, 其中 α 控制初始决策值对相邻节点颜色空间距离的调整幅度, β 为随机采样方式下两个像素点颜色差的方差,试验中 α 取 0.5, β 取值与 Grab cut^[7]方法一致。

接下来,考虑函数的第2部分。3维空间中,使用 k 最近邻方法构造节点间的连接关系,每个节点的特征使用空间位置和颜色特征,距离描述为高斯核映射距离而权重为归一化的核距离。为保证存在连接边的两个节点能在至少同一图像中存在投影点,在建立的 k 最近邻图后要滤除没有公共投影图像的边。本文使用鉴别直推式学习模型^[13] (discriminative transductive learning model)求解3

维空间分割问题。为保证数学描述的一致性,在函数中 $\forall x^{(3D)} \in [0,1]$, $p^{(3D)} = 2$ 的条件下,对 E_2 推导如下:

$$\left. \begin{aligned} E_2 = E_p(\mathbf{x}^{(3D)}) &= \sum_{e_{ij}^{(3D)} \in \mathcal{E}^{(3D)}} \omega_{ij}^{(3D)} |x_i^{(3D)} - x_j^{(3D)}|^2 \\ &= \sum_{e_{ij}^{(3D)} \in \mathcal{E}^{(3D)}} \omega_{ij}^{(3D)} (x_i^{(3D)} - x_j^{(3D)})^2 + \sum_{\mathcal{O}_M^{(3D)}} \omega_i^{(3D)} (x_i^{(3D)})^2 \\ &+ \sum_{\mathcal{B}_M^{(3D)}} \omega_i^{(3D)} |x_i^{(3D)} - 1|^2 \\ &= \mathbf{x}^{(3D)\top} \mathbf{L} \mathbf{x}^{(3D)} + \sum_{i \in \mathcal{O} \cup \mathcal{B}} \omega_{ij}^{(3D)} (\mathbf{y} - \mathbf{x})^2 \\ &= \mathbf{x}^{(3D)\top} \mathbf{L} \mathbf{x}^{(3D)} + (\mathbf{y}^{(3D)} - \mathbf{x}^{(3D)})^\top \mathbf{C} (\mathbf{y}^{(3D)} \\ &- \mathbf{x}^{(3D)}) \\ \text{s.t. } x(\mathcal{O}_M^{(3D)}) &= 1, x(\mathcal{B}_M^{(3D)}) = 0, y_i^{(3D)} = 1 \text{ if } x_i^{(3D)} \\ &\geq \frac{1}{2}, 0 \text{ if } x_i^{(3D)} < \frac{1}{2} \end{aligned} \right\} (3)$$

其中 $\mathbf{y}^{(3D)}$ 是种子点点集的标签向量表示, \mathbf{L} 矩阵为图谱聚类中的拉普拉斯矩阵, $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}$, $\mathbf{W}_{ij}^{(3D)} = k^{(3D)}(p_i, p_j)$, $\mathbf{D}_{ii}^{(3D)} = d^{(3D)}(p_i)$, \mathbf{C} 为对角矩阵, 对角元素为种子点错误分类时的代价权重。此时,对 E_2 的最小化问题转换为半监督归一化割(semi-supervised normalized cut)求解问题,本文采用文献[20]提出的直推式谱图(spectral graph transducer)方法进行求解。

通过上面两步,单独最小化 E_1 和 E_2 均已完成,接下采用迭代优化方法最小化 $E_p(\mathbf{x})$ 。 $E_p(\mathbf{x})$ 中的约束条件是3维点标签和投影后对应图像像素点的标签一致性约束。其中 $y_{\mathbf{x}_M}^{(3D)} = y_{\mathbf{x}_M}^{(k)}$ 指对应种子点的标记必须一致, $y_{u_i}^{(3D)} = y_{u_i}^{(k)}$ 指3维点 p_i 投影到图像 k 上,像素坐标为 $u_i^{(k)}$, 那么像素上的标记和3维点标记最终必须一致。迭代前,待分割图像均使用快速无监督图割方法^[21]对图像进行初始分割,将图像划分为 m 个不重叠区域 $R^{(k)} = \{r_j^{(k)}\}, j = 0, \dots, m$, 而投影点的决策值 $x_{u_{i,t}}^{(k)}$ 将传播到属于该区域的所有像素点,在得到第 K 次分割结果 I^K 后,使用新的 $\mathcal{O}_i^{I^K}, \mathcal{B}_i^{I^K}$ 重新训练全局高斯混合模型 $\text{GMM}_{i,\mathcal{O}}^G$, $\text{GMM}_{i,\mathcal{B}}^G$ 。3维空间标记的更新采用多幅图像中的对应标记投票完成,而前后两次迭代中标记不一致的3维点将被过滤掉,不作为种子点考虑。上述迭代过程属于贪婪算法,在初始值良好, E_1 和 E_2 最小化解分别能达到全局最优情况下,对本文非凸优化问题是行之有效的,实验中将给予更多说明。

4 实验及结果

本文实验使用两组图像序列,首先基于文献[2]

方法得到3维重建点。图像采集约束为均匀光照下,目标表面可近似为朗伯表面。第1组图像数目为30幅;第2组图像数目为9幅。这两组实验用图,图像数目多而且复杂度高,手工逐一分割工作量很大。两组实验所用参数均相同,分割时图像尺寸统一缩放至 320×240 。以下参数可根据不同的图像序列作调整,其中前景背景可分别使用3-6个混合高斯,粗分割中不重叠区域 m 大小可在20到600间调整,粗分割若分割区域较大,容易导致传播过大,而如果分割后区域过小,又会增加算法迭代次数。本文对用户交互点使用粗分割与最近邻结合的传播原则,控制用户交互的种子点传播的范围,这种方式稳定可靠,但图像复杂情况下会增加交互次数。

图2为初始选择1幅图后,使用Grab cut方法交互式分割的结果,图2(a1),(a2)为用户选择的初始图像;图2(b1),(b2)为节点间的连接权重,灰度值越大代表权重越大;图2(c1),(c2)为背景目标分割结果及轮廓带,分别用黑白灰表示,轮廓带宽度设置为7;图2(d1,d2)为在该视图下可见的3维点向其投影,这样就可以由图像分割结果决定3维空间种子点。图2(c1)中的矩形框为用户交互的包含目标的边框,曲线为用户交互提供的前景或背景种子点。

图3,图4分别为一组图像及3维空间的分割结果,图3,图4中(a1)为初始的3维空间种子点(方框内白色为前景种子,方框外白色为背景种子,灰色为未标记点);图3,图4中(a2)为3维空间的分割结果(白色点为预测的前景点,灰色点为预测的背景点);图3,图4中(b1),(b2)是不同视角下图像的无监督粗分割结果,不同区域用不同颜色表示,该结果用于图像空间种子点的传播;图3,图4中(c1),(c2)为不同视角下的图像自动分割结果(背景区域被淡化)。

算法流程中过滤操作主要体现在将投影点落在轮廓带内的点滤除,不纳入种子点,这是因为目前基于运动恢复结构的方法首先都需要计算图像中的特征点,而特征点一般都出现在图像边缘角点等位置,因此恢复出的3维点投影后也出现在这些位置,因此由于投影矩阵或交互式分割结果的微小误差,就会导致这些位置设置种子点,其标签错误的情况,而种子点的好坏严重影响分割效果,因此过滤操作十分重要。传播操作中,首先采用粗分割方法如快速图割法、均值偏移法等将图像划分为一组图像块,将满足条件的种子点传播到其所在的图像块中。这种方法避免了样本点不足,过分割产生的孤立点等问题。投票操作的思想是单独一幅图像的分割结果存在错误,而多幅图像分割结果进行投票决定3维空间种子点,有效阻止了错误累积问题。

本文算法使用C++编写,运行硬件环境为Intel i7 860, 3 G内存,做到实时交互。第1组图迭代9次后收敛,第2组图迭代5次后收敛。分割结果的精度指标采用平均误判率,即:每幅图中错误标记像素的个数除以总像素个数,最后所有图取平均。基准由Photo Shop逐一手动交互分割目标得到。第1组图像平均误判率为2.7%,第2组图像平均误判率为4.1%,据我们了解,目前还没有文献提供该问题的实验数据,因此我们只给出了与人工基准数据的比较。由于本算法提供用户选择性交互步骤,不限定交互次数的话,误判率是可以接近0.0%。(交互次数未知情况下,3维信息仍可提供比Photo Shop多的约束,减少大量不必要交互)。上述结果是在限定用户交互次数不超过平均每幅图2次下得到的。实验中每幅图像独立使用Grab cut方法^[7]并限定每幅图交互次数为2次,得到的平均误判率分

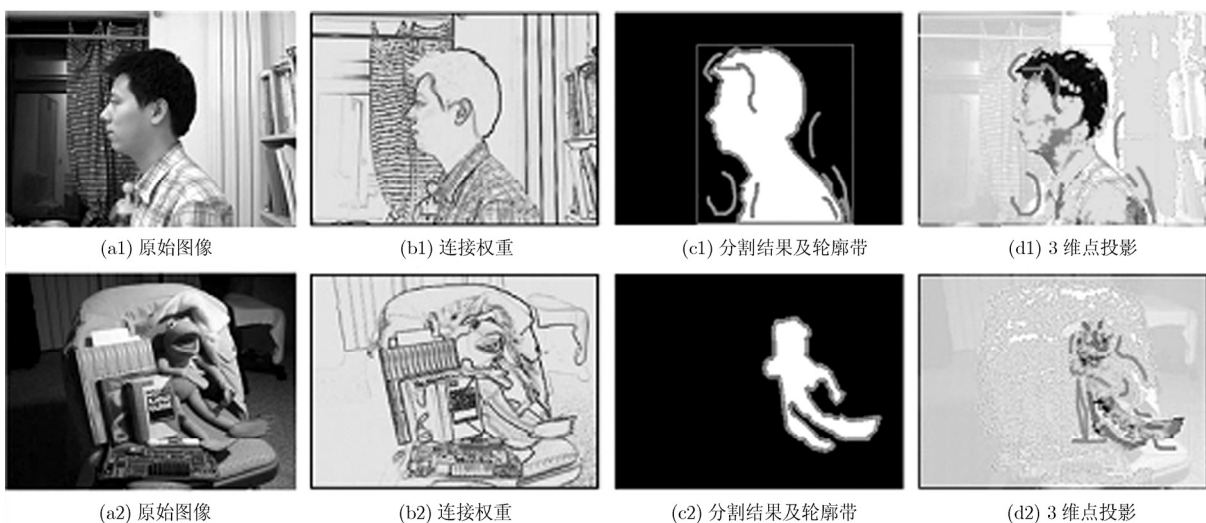


图2 初始交互式分割



图 3 迭代式协同分割结果 1

图 4 迭代式协同分割结果 2

别为 5.9%和 9.3%。本文算法在精度和效率上均有较大提高, 还需要指出的是, 本文方法除给出每幅图像中的目标分割结果外, 同时得到 3 维空间中目标的表面部分重建点, 而这些结果可直接用于基于图像的目标建模研究, 提高目标重建的精度。

5 结束语

本文重点研究多视角拍摄刚性或静态目标图像后, 在用户交互情况下快速完成多图像目标分割这个问题。本文提出的直推式迭代协同分割方法, 实验证明了算法的有效性并给出了精度评价, 但目前, 算法对于目标颜色与大面积邻近背景颜色基本相同的图像序列, 结果较差, 主要原因是这类图像序列重建结果中 3 维点分布很不均匀, 背景点样本不足, 此时基于颜色的混合高斯模型, 导致大量后验概率错误, 因此, 采用局部区域纹理描述代替点颜色描述是本文接下来要改进的方向。

参考文献

[1] Batra D, Kowdle A, Parikh D, Luo Jiebo, and Chen Tsuhan. iCoseg: interactive co-segmentation with intelligent scribble guidance[C]. Computer Vision and Pattern Recognition (CVPR '10), San Francisco, USA, 13-18 June., 2010: 1-8.

[2] Agarwal S, Snavely N, Simon I, Seitz S M, and Szeliski R. Building rome in a day[C]. 12th IEEE Int'l Conf. Computer

Vision (ICCV'09), Kyoto, Japan, Sept 29-Oct 2, 2009: 72-79.

[3] Furukawa Y and Ponce J. Carved visual hulls for image-based modeling[J]. *International Journal of Computer Vision*, 2009, 81(1): 53-67.

[4] Liu Ye-bin, Dai Qiong-hai, and Xu Wen-li. Continuous depth estimation for multi-view stereo[C]. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR'09), Florida, USA, June 2009: 2121-2128.

[5] Boykov Y and Jolly M P. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images[C]. ICCV'01, Vancouver, British Columbia, Canada, 7-14 July, 2001: 105-112.

[6] Grady L. Random walks for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(11): 1768-1783.

[7] Rother C, Kolmogorov V, and Blake A. Grabcut: interactive foreground extraction using iterated graph cuts[C]. International conference on computer graphics and interactive techniques (SIGGRAPH'04), Los Angeles, USA, 10-12 August, 2004: 309-314.

[8] Duchenne O, Audibert J Y, Keriven R, Ponce J, and Segonne F. Segmentation by transduction[C]. CVPR'08, Anchorage, Alaska, USA, 23-28 June, 2008: 1-8.

[9] Meila M and Shi Jian-bo. A random walks view of spectral segmentation[C]. International Conference on AI and

- Statistics, Key West, Florida, USA, 4-7 January, 2001: 1-8.
- [10] Shi Jian-bo and Malik J. Normalized cuts and image segmentation[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905.
- [11] Couprie C, Grady L, Najman L, and Talbot H. Power Watersheds: a new image segmentation framework extending graph cuts, random walker and optimal spanning forest[C]. ICCV'09, Kyoto, Japan, Sept. 29-Oct. 2, 2009: 731-738.
- [12] Cui Jing-yu, Yang Qiong, Wen Fang, Wu Qi-ying, Zhang Chang-shui, Van Cool L, and Tang Xiao-ou. Transductive object cutout[C]. CVPR'08, Anchorage, Alaska, USA, 23-28 June, 2008: 1-8.
- [13] Zhu Xiao-jing and Goldberg A B. Introduction to Semi-supervised Learning[M]. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 3:1, USA. Morgan & Claypool Publishers, 2009: 1-130.
- [14] Yin Pei, Criminisi A, Winn J, and Essa I. Bilayer segmentation of webcam videos using tree-based classifiers[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2010, 33(1): 30-42.
- [15] Kolmogorov V, Criminisi A, Blake A, Cross G, and Rother C. Bi-layer segmentation of binocular stereo video[C]. CVPR'05, San Diego, CA, USA, 20-26 June, 2005: 407-414.
- [16] Quan Long, Wang Jing-dong, Tan Ping, and Yuan Lu. Image-based modeling by joint segmentation[J]. *International Journal of Computer Vision*, 2007, 75(1): 135-150.
- [17] Xiao Jianxiong, Wang Jingdong, Tan Ping, and Quan Long. Joint affinity propagation for multiple view segmentation[C]. ICCV'07, Rio de Janeiro, Brazil, 14-20 October, 2007: 43-51.
- [18] Komodakis N, Paragios N, and Tzirita G. MRF optimization via dual decomposition: message-passing revisited[C]. ICCV'07, Rio de Janeiro, Brazil, 14-20 October, 2007: 1-8.
- [19] Yang Jian-chao, Yu Kai, and Huang T. Supervised translation-invariant sparse coding[C]. CVPR'10, San Francisco, USA, 13-18 June, 2010: 1-8.
- [20] Joachim T. Transductive learning via spectral graph partitioning[C]. *International Conference on Machine Learning (ICML)*, Washington, DC, USA, 21-24 August, 2003: 290-297.
- [21] Felzenszwalb P F and Huttenlocher D P. Efficient Graph-based image segmentation[J]. *International Journal of Computer Vision*, 2004, 59(2): 167-181.
- 朱云峰: 男, 1981年生, 博士生, 研究方向为图像分割、基于图像的渲染。
- 章毓晋: 男, 1954年生, 教授, 博士生导师, 研究方向为图像工程。