

基于局部特征的多目标图像分割算法

王琳^{1,2}, 刘强^{1,2}

¹天津大学微电子学院, 天津 300072;

²天津市成像与感知微电子技术重点实验室, 天津 300072

摘要 近年来,智能机器人技术逐步成熟,以目标识别为代表的机器视觉技术是其核心。现有目标识别算法通常先根据颜色信息分割场景,再提取特征以识别目标。但对于颜色信息比较复杂的场景,往往存在过分割现象,影响后续目标识别。针对这一问题,提出一种基于局部特征的多目标图像分割算法。该算法使用双目摄像头采集场景图像,对场景图像进行预处理,同时通过立体匹配得到场景的深度信息;结合深度信息确定目标区域;设计动态阈值的尺度不变特征变换(SIFT)算法以提取目标区域的局部特征,将局部特征转化为特征约束;基于区域约束、特征约束和空间信息组成的特征向量进行聚类分割,得到最终分割结果,同时实现对每个目标区域的识别。实验结果表明,在颜色特征复杂的场景中,本文算法的整体误差率小于10%,与已有算法相比降低了15%以上。

关键词 图像处理; 多目标分割; 尺度不变特征; 均值漂移

中图分类号 TP391.4 文献标识码 A

doi: 10.3788/LOP55.061002

A Multi-Object Image Segmentation Algorithm Based on Local Features

Wang Lin^{1,2}, Liu Qiang^{1,2}

¹School of Microelectronics, Tianjin University, Tianjin 300072, China;

²Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology, Tianjin 300072, China

Abstract Intelligent robot has recently matured in industry, whose core technology is machine vision, especially object recognition. In existing object recognition methods, scenes are segmented based on color, and features are then extracted to recognize objects. However, over segmentation exists for scenes with complex color features, which influences subsequent object recognition process. To deal with the over segmentation problem, a multi-object image segmentation algorithm based on local features is proposed, which uses binocular camera to collect scene images. Firstly, the scene image is preprocessed. The depth information of the scene is then obtained by stereo matching, and is used to determine the target area. Secondly, the local features of the target region are extracted by a scale-invariant feature transform (SIFT) algorithm with dynamic threshold, and the local features are then transformed into feature constraints. Finally, the feature vectors, which are based on region constraint, feature constraint and spatial information, are used for clustering segmentation to obtain the final segmentation result. Simultaneously, each object region is recognized. The experiment results show that the overall error rate of the proposed algorithm is less than 10% for a scene with complex color features, and is reduced by 15% compared with those of the existing algorithms.

Key words image processing; multiple-object segmentation; scale invariant feature; mean shift

OCIS codes 100.2000; 100.3008

1 引言

近年来,视觉智能机器人逐渐在物流、制造等行

业中广泛应用^[1]。视觉机器人的核心技术之一是目标识别。现有目标识别算法一般将图像分割作为前置算法,再提取特征、识别目标,所以分割结果的好

收稿日期: 2017-11-02; 收到修改稿日期: 2017-11-24

基金项目: 国家自然科学基金(61574099)

作者简介: 王琳(1992—),男,硕士研究生,主要从事图像处理、硬件加速方面的研究。E-mail: devil@tju.edu.cn

导师简介: 刘强(1978—),男,博士,副教授,博士生导师,主要从事图像处理、硬件安全方面的研究。

E-mail: qiangliu@tju.edu.cn(通信联系人)

坏直接影响识别的效果^[2]。均值漂移算法及其改进算法被广泛用于图像分割^[2]，但单一地利用均值漂移算法基于彩色信息进行图像分割并不能取得预期效果^[3]。文献[4]在均值漂移聚类过程中引入边缘敏感度的概念，并融合深度信息对场景进行分割，使其在分割颜色接近且相互遮挡的物体时具有较好的效果，但仍存在过分割的问题。文献[5]提出在均值漂移算法的基础上，通过区域显著性特征融合过分割的初始结果，以得到较理想的最终分割结果，但当场景中的目标过小时，容易出现分割错误的情况。这些方法主要基于颜色特征进行分割，在处理拥有复杂颜色信息的场景时效果不佳。

最近，全卷积神经网络(FCN)^[6]被成功用于图像分割^[7]。利用卷积神经网络的学习能力，即使使用未经处理的原始图像作为输入，也可以得到较准确的结果。然而，FCN算法也存在一些缺点：一是需要大量的数据进行训练，以获得期望的图像分割结果；二是分割结果不够精细，尤其是在边缘细节方面，分割图像在边缘上会有一定程度的模糊^[7]。

针对以上问题，本文基于传统图像处理方法，提出一种基于局部特征的多目标图像分割算法。该算

法的核心思想是先提取图像的局部特征，然后将局部特征与深度信息、位置信息结合生成新的特征向量，最后使用均值漂移算法对特征进行聚类分割，从而在拥有复杂颜色信息的场景中可以得到较好的分割效果。算法的创新点主要包括：1) 结合局部特征、深度信息和位置信息进行图像分割，改善了复杂颜色信息场景中的分割效果；2) 采用自适应背景差分算法确定目标主体区域，减小计算量，提高实时性；3) 针对多个目标模板，改进尺度不变特征变换(SIFT)算法，降低误匹配率。

2 多目标分割算法

本文算法流程如图1所示。该算法使用双目摄像头采集场景图像，首先对场景图像进行预处理，同时通过立体匹配得到场景的深度信息。结合深度信息确定目标区域，获得区域约束。针对多个目标模板，采用基于动态阈值的SIFT算法提取目标区域的局部特征，降低误匹配率，将局部特征转化为特征约束，提高后续分割的精确度。最后，基于区域约束、特征约束和空间信息组成的特征向量进行均值漂移聚类分割，得到最终分割结果的同时实现对每个目标区域的识别。

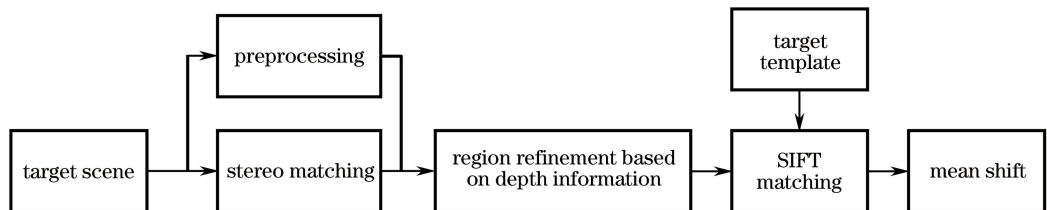


图1 本文算法流程图

Fig. 1 Flow chart of the proposed algorithm

2.1 预处理

该步骤对应两种场景：1) 可以预先获取背景图像，如仓库分拣货物场景，使用自适应背景差分算法，减小后续步骤的计算量；2) 无法预先获取背景图像，则可以省略该步骤，在后续步骤中分割背景区域，但计算量有所增加。

在可以预先获取背景图像的情况下，借鉴运动目标检测中的背景差分法，本文算法将含有目标的场景图像与背景图像进行差分处理，初步确定目标范围。具体步骤如下。

定义 $B(x, y)$ 为背景图像， $F(x, y)$ 为加入目标后的场景图像， $D(x, y)$ 为差分图像， $I(x, y)$ 为二值化图像。

将图像进行 RGB 三通道分离，对各通道分别进

行背景差分，对每个像素选取各通道差值的最大值，进行阈值二值化：

$$D_i(x, y) = |F_i(x, y) - B_i(x, y)|, \quad (1)$$

$$I(x, y) = \begin{cases} 1, & \max[D_i(x, y)] \geq \theta^B \\ 0, & \max[D_i(x, y)] < \theta^B \end{cases}, \quad (2)$$

式中 $i=r, g, b$ 表示三个颜色分量； $F_i(x, y)$ 为场景图像 $F(x, y)$ 分离出的第 i 通道的图像； $B_i(x, y)$ 为背景图像 $B(x, y)$ 分离出的第 i 通道的图像； $D_i(x, y)$ 为第 i 通道的差分图像； θ^B 为设定的二值化阈值。

理想情况下，二值化图像中目标区域的值为 1，背景区域的值为 0。但在实际图像采集中，摄像头会根据视场的变化自动调整曝光参数，以得到亮度正常的图像。当一个目标放入摄像头视场内时，采

集图像的背景亮度会发生改变,在 RGB 颜色空间中表现为每个像素点三个通道的值以一定比例变化。因此,在不同的场景图像中使用相同的固定阈值 θ^B ,在二值化过程会将背景部分误认为目标。

针对这一问题,采用自适应法确定二值化阈值。对不同图像,通过一个迭代算法确定最优阈值。目的是使背景图像和场景图像背景区域的平均亮度接近,消除背景亮度变化的影响。自适应阈值确定法的具体步骤如下。

首先,定义 R' 为当前二值化阈值下 $\mathbf{I}(x, y) = 0$ 的背景区域, d 为背景区域平均差值,则 d 的计算公式为

$$d = \frac{\sum_{x,y \in R'} \mathbf{D}(x, y)}{|R'|} \quad (3)$$

对于 θ_i^B , 计算 $\mathbf{I}(x, y)$, 然后求得 d_i 。

$$\theta_{i+1}^B = \theta_i^B + d_i - d_{i-1}, \quad (4)$$

利用(4)式更新 θ_i^B , 直至 $d_{i+1} = d_i$, 迭代结束。此时,背景图像和场景图像背景区域的平均亮度相同。

应用背景差分算法进行二值化的实例如图 2 所示。图 2(c) 为使用固定阈值二值化的结果,部分墙壁、桌面等背景区域被识别为目标区域。图 2(d) 为使用自适应阈值的二值化结果,识别的目标区域与实际目标区域基本吻合。

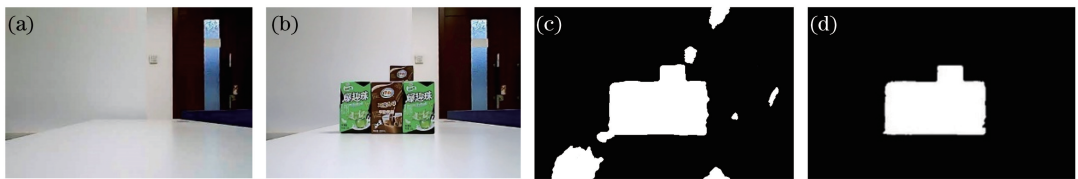


图 2 (a)背景图像;(b)场景图像;(c)固定阈值结果;(d)自适应阈值结果

Fig. 2 (a) Background image; (b) scene image; (c) result with fixed threshold; (d) result with adaptive threshold

2.2 深度信息提取

在深度信息的获取中,双目立体视觉具有计算效率高、精度适中、系统结构简单、成本低等特点,逐步成为计算机视觉领域的研究热点^[8]。其中图像的立体匹配是立体视觉中的核心技术^[9]。由两个摄像头组成双目视觉系统,通过立体匹配算法获取场景深度信息的方法成为首选。

通过双目摄像头对同一场景采集左、右两个视角的图像,再使用立体匹配算法计算得到图像的深度信息。之后对立体匹配得到的图像进行平滑、填补空洞等处理,得到深度图像 $\mathbf{I}_D(x, y)$ 。

2.3 基于深度信息的目标区域细化

基于深度信息对目标区域进行进一步细化。针对可以获得背景图像的场景,细化后的图像 $\mathbf{V}(x, y)$ 为

$$\mathbf{V}(x, y) = \begin{cases} 0, & \mathbf{I}(x, y) = 0 \\ \mathbf{I}_D(x, y), & \mathbf{I}(x, y) = 1 \end{cases} \quad (5)$$

针对无法获得背景图像的场景,细化后的图像 $\mathbf{V}(x, y)$ 为

$$\mathbf{V}(x, y) = \mathbf{I}_D(x, y). \quad (6)$$

对细化后的图像 $\mathbf{V}(x, y)$ 中值不为 0 的区域进行连通域分析,按照四邻接规则将值相同的像素标记在同一连通域中。记获得的 n 个连通域分别为 $O_k, k = 1, 2, \dots, n$, 进一步得到区域约束 $\mathbf{L}(x, y)$ 为

$$\mathbf{L}(x, y) = \begin{cases} k, & x, y \in O_k \\ 0, & x, y \notin O_k \end{cases} \quad (7)$$

2.4 局部特征提取及多模板匹配

SIFT 算法是一种基于尺度空间理论的局部特征匹配算法^[10], 拥有对尺度、光照、旋转等各种变换的不变性、多量性和独特性^[11], 提取的局部特征具有较高的辨识度^[12], 使用这种局部特征进行图像匹配可以实现对目标的识别。然而,当不同目标的特征相近时,容易造成误匹配。

针对以上问题,基于获得的目标区域,对 SIFT 算法的特征提取及匹配过程进行改进。匹配结果用于后续聚类分割。

2.4.1 特征模板库

SIFT 算法可以分为特征提取与特征匹配两部分。在特征提取步骤中,需要构建高斯金字塔,并在高斯金字塔中检测、定位特征点。该步骤计算量较大,所需时间较长^[13]。提取多个目标模板的特征点无疑会使时间消耗成倍增长,算法的实时性下降。

为减少特征提取过程的时间消耗,本文算法预先提取每个模板的 SIFT 特征,并使用这些特征建立模板特征库。在实际运用时,仅需要提取一次场景图像的特征点,这将大大降低该过程的时间消耗。从特征库读取模板特征耗时 0.45 s, 而直接从模板图像中提取模板特征需要耗时 1.76 s。

2.4.2 自适应阈值匹配

SIFT 算法需要设置一个阈值 θ^s 。当目标特征点的特征矢量与模板特征点的特征矢量之间的欧氏距离小于 θ^s 时,两个特征点匹配。

如果不同模板之间存在相近特征,则一个特征点会与多个模板匹配。这种情况会使后续聚类分割出现困难。为此,先对模板图像库中的模板两两进行 SIFT 匹配,为每个模板确定一个阈值:

$$\theta_i^s = \min(\theta_{ij}^s - 1), \forall j \neq i, \quad (8)$$

式中 θ_i^s 为模板 i 的阈值, θ_{ij}^s 为模板 i 和模板 j 可以匹配的最小欧氏距离。在这一阈值下,每个模板都

表 1 自适应阈值与固定阈值匹配点比较

Table 1 Comparison of matching points between adaptive threshold and fixed threshold

Parameter	Adaptive threshold	Fixed threshold		
		$\theta^s = 150$	$\theta^s = 200$	$\theta^s = 250$
Number of feature points	50	40	55	89
Number of mismatching feature points	1	2	3	13
Mismatching rate /%	2.0	5.0	5.5	14.6

2.5 均值漂移聚类分割

将得到的每个像素点的区域约束 $\mathbf{L}(x, y)$ 和特征约束 $\mathbf{T}(x, y)$, 与该像素点坐标 (x, y) 结合, 组成新的特征向量 $[x, y, \mathbf{L}(x, y), \mathbf{T}(x, y)]$ 。该特征向量与基于颜色信息进行分割的特征向量 $[x, y, \mathbf{V}_{\text{RGB}}(x, y)]$ 相比, 减小了目标本身复杂颜色特征的影响, 可以避免过分割现象的发生。基于新的特征向量, 使用均值漂移算法对特征点进行聚类、分割。主要步骤如下:

Input: $\mathbf{B}(x, y), \mathbf{F}(x, y)$

Output: $\text{Result}(x, y)$

//preprocessing

for: $i < -1$ to height

for: $j < -1$ to width

$D(i, j) = \max(\mathbf{F}_i(i, j) - \mathbf{B}_i(i, j))$ //background subtraction

do

$\theta_{i+1}^b < -\theta_i^b$

$\mathbf{R}'_{i+1} < -\mathbf{R}'_i$

$d_{i+1} < -d_i$

while $d_{i+1} = d_i$ //calculation threshold

$\mathbf{I}(x, y) < -\text{Binaryzation}(D(x, y), \theta_i^b)$ //binaryzation

//depth image extraction

$\text{Depth}(x, y) < -\text{StereoMacthing}(\mathbf{F}_l(x, y), \mathbf{F}_r(x, y))$ //stereo matching

//region refinement based on depth information

for: $i < -1$ to height

for: $j < -1$ to width

if $\mathbf{I}(i, j) = 0 \quad \mathbf{V}(i, j) = 0$

不会与其他模板匹配,可以减少一个目标与多个模板匹配的情况。如果某个特征点仍与多个模板匹配,则删除该特征点。由于只是删除个别特征点,不会对后续目标识别产生负面影响。

每个特征点 T_i 均与一个模板 M_j 匹配,则图像的局部特征转化为特征约束:

$$\mathbf{T}(x, y) = \begin{cases} M_j, & (x, y) \in T_i \\ 0, & (x, y) \notin T_i \end{cases} \quad (9)$$

表 1 为(8)式计算得到的阈值与固定阈值的特征点比较。可以看出通过(8)式得到的阈值可以获得更多的有效特征点,且误匹配率更低。

1) 根据特征向量中的 $\mathbf{T}(x, y), \mathbf{L}(x, y)$ 特征将特征点划分为不同组;

2) 对每组特征点基于其特征向量中的空间特征进行均值漂移聚类;

3) 将聚类后的特征点重新分组;

4) 根据区域约束中的深度信息,得到最终分割结果。

本文算法整体流程的伪代码如下。

```
else V(i,j)=Depth(i,j) //region refinement
```

```
L(x,y)← Connected (V(x,y)) //connected domain analysis
```

```
//SIFT matching
```

```
Ttemp ← sift_point_Extract(L(x,y),F(x,y)) //SIFT local feature extraction
```

```
for i← 1 to template_num
```

```
    T(x,y)← sift_match(Ttemp,  $\theta_i^s$ ) //SIFT local feature matching
```

```
//Mean Shift
```

```
Feature_vector ← [x,y,L(x,y),T(x,y)]
```

```
Result(x,y) ← MeanShift([x,y,L(x,y),T(x,y)])
```

3 实验及结果分析

为验证算法的有效性,搭建了图 3 所示 8 个含有多种颜色目标的场景,图像的分辨率为 320 pixel×240 pixel。实验分为三部分:1) 评估本文 SIFT 算法和原始 SIFT 算法的误匹配率;2) 应

用本文算法、文献[5]算法和 FCN 算法^[6]对 8 幅图像分别进行分割,比较分割结果;3) 分析本文算法、文献[5]算法和 FCN 算法^[6]的实时性。实验环境: Intel[®] Xeon[™] CPU E5-2630 v4 @ 2.20 GHz, NVIDIA[®] Tesla[™] GPU K40,配有 32 GB 内存;安装有 Caffe 计算框架及 OpenCV 2.4.10 函数库。



图 3 8 种实验场景

Fig. 3 Eight experimental scenarios

为了验证本文算法在不同情况下的性能,搭建图 3 所示 8 个场景,对应验证情况如下:

- 1) 多个目标区域在图像中不相连(场景 1、7、8);
- 2) 多个目标区域在图像中相连(场景 2~6);
- 3) 后排目标遮挡较多(场景 4~6);
- 4) 背景较为复杂(场景 7、8)。

3.1 局部特征提取及匹配结果分析

表 2 为 8 种场景下原始 SIFT 算法与本文针对多模板匹配优化后的 SIFT 匹配结果的对比。从实验结果可以看出,由于使用自适应阈值,本文算法使误匹配特征点数量减少。与直接使用原始 SIFT 算法相比,误匹配率大大降低。

3.2 分割结果分析

图 4 和图 5 分别为针对场景 5 和场景 8 使用文献[5]算法、FCN 算法^[6]、本文算法的分割结果对比。从图 4 可以看出,场景 5 的背景较为简单,多个目标在图像的同区域内,从分割结果看,本文算法

表 2 本文 SIFT 算法与原始 SIFT 算法误匹配率比较
Table 2 Comparison of mismatching rate between the proposed SIFT algorithm and the original SIFT algorithm

Scene	Mismatching rate /%	
	Original SIFT algorithm	Proposed SIFT algorithm
1	18.9	1.5
2	10.0	0
3	9.3	1.7
4	9.7	0
5	14.3	2.0
6	10.4	1.9
7	22.5	1.2
8	19.8	2.7
Average	13.2	1.4

可以分割出场景中的目标,与 FCN 算法^[6]得到的分割结果相近,均优于文献[5]算法的分割结果;而场景 8 的背景较为复杂,多个目标分散在不同区域,从图 5 可以看出,本文算法同样适用于背景相对复杂且有多个分散目标的场景。

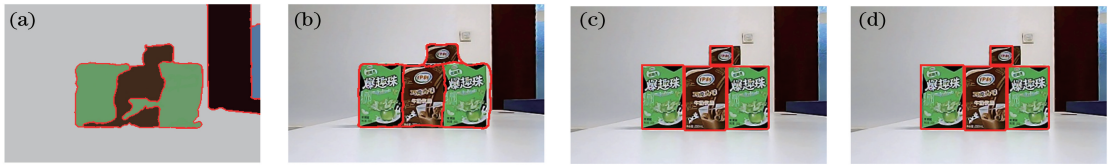


图4 针对场景5使用不同算法得到的分割结果。(a)文献[5]算法;
(b) FCN 算法;(c)本文算法(含预处理);(d)本文算法(不含预处理)

Fig. 4 Segmentation results obtained by different algorithms for the scene No. 5. (a) Algorithm in Ref. [5]; (b) FCN algorithm; (c) proposed algorithm with preprocessing; (d) proposed algorithm without preprocessing

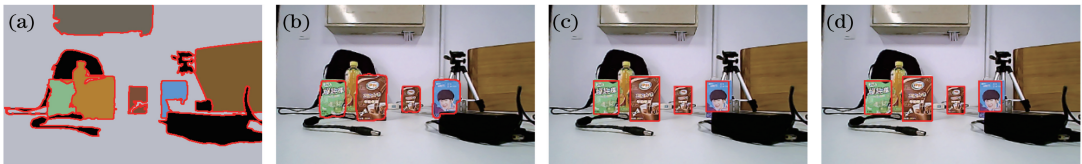


图5 针对场景8使用不同算法得到的分割结果。(a)文献[5]算法;
(b) FCN 算法;(c)本文算法(含预处理);(d)本文算法(不含预处理)

Fig. 5 Segmentation results obtained by different algorithms for the scene No. 8. (a) Algorithm in Ref. [5]; (b) FCN algorithm; (c) proposed algorithm with preprocessing; (d) proposed algorithm without preprocessing

图像分割效果可以用图像分割误差指标来衡量^[14],包括过分割率(R_O)、欠分割率(R_U)和整体错误率(R_E)。

$$R_O = \frac{O_s}{R_s + U_s}, \quad (10)$$

$$R_U = \frac{U_s}{R_s + U_s}, \quad (11)$$

$$R_E = \frac{O_s + U_s}{R_s}, \quad (12)$$

式中 R_s 为手动分割得到的目标像素数目; O_s 为本应该包含在分割结果中,实际却不在分割结果中的像素数目; U_s 为本不应该包含在分割结果中,实际却在分割结果中的像素数目。本文算法与 FCN 算法、文献[5]算法的分割误差比较如表3所示。

表3 本文算法与 FCN 算法、文献[5]算法的分割误差比较

Table 3 Comparison of segmentation errors among FCN algorithm, the algorithm in Ref. [5] and the proposed algorithms

Scene		1	2	3	4	5	6	7	8	Average
Proposed algorithm (with preprocessing)	R_O /%	3.40	4.38	7.49	6.47	9.84	9.86	9.18	9.28	7.49
	R_U /%	0	0	0	0	0	0	0	0	0
	R_E /%	3.40	4.38	7.49	6.47	9.84	9.86	9.18	9.28	7.49
Proposed algorithm (without preprocessing)	R_O /%	3.40	4.31	7.49	6.17	8.95	9.83	7.87	8.09	7.01
	R_U /%	0	1.70	0	0.85	1.29	0.26	2.65	1.20	0.99
	R_E /%	3.40	6.11	7.49	7.08	10.37	10.12	10.80	9.40	8.10
FCN algorithm	R_O /%	2.42	1.17	3.53	4.04	5.94	5.72	7.70	4.70	4.40
	R_U /%	1.12	4.36	2.95	2.08	3.34	2.54	1.82	3.81	2.74
	R_E /%	3.58	6.79	6.68	6.25	9.48	8.47	9.69	8.85	7.47
Algorithm in Ref. [5]	R_O /%	18.19	25.00	20.39	23.26	14.26	31.88	14.11	11.77	19.86
	R_U /%	0	23.12	15.79	13.04	10.20	26.67	3.23	15.44	13.44
	R_E /%	18.19	62.59	42.96	41.74	27.25	79.83	17.92	32.18	40.33

对于拥有复杂颜色特征的背景和目标,文献[5]算法基于彩色和空间特征进行均值漂移分割后会有严重的过分割现象,且无法区分背景和应识别的目标[图4(a)和图5(a)]。主要原因是场景中的目标在整幅图像中所占比例过小,且不同物体之间颜色

相近,造成错误的分割结果。因此文献[5]算法的整体错误率偏高,不适用于本文的应用环境。

FCN 算法^[6]的分割结果较为准确,但分割结果的边缘细节不够理想[图4(b)和图5(b)]。对于图像中目标区域不相连或相连较少的场景,整体错误

率较低;随着图像中目标区域连接的增加,欠分割率、过分割率和整体错误率均不断增大。

当可以预先获得背景时,本文算法含有预处理步骤,其最终分割结果如图 4(c)和图 5(c)所示,每个红色方框包围的区域为最终目标区域,其余为背景区域。相应地,当无法预先获取背景时,本文算法不含预处理步骤,其最终分割结果如图 4(d)和图 5(d)所示。从表 3 误差分析结果可以看出,含有预处理步骤的平均整体错误率为 7.49%,不含预处理步骤的平均整体错误率为 8.10%。对于没有遮挡的前排目标,分割结果较为理想;后排目标在非遮挡情况下也可以得到较为满意的结果;但后排目标在遮挡情况下的误差较大。因此在遮挡较多的场景中,本文算法的整体错误率较高。

本文算法的过分割率高于 FCN 算法^[6]的过分割率,而欠分割率优于 FCN 算法^[6]。从整体错误率来看,本文算法分割结果与 FCN 算法^[6]分割结果相近。在无遮挡或遮挡较少的场景中,FCN 算法^[6]的分割边缘使整体错误率较高,本文算法分割结果优于 FCN 算法^[6];但在遮挡较多的场景中,本文算法对遮挡目标的分割误差较大,FCN 算法^[6]分割结果优于本文算法。

3.3 时间复杂度及实时性分析

本文算法中基于颜色信息的背景差分步骤的时间复杂度为 $O(T_b mn)$, T_b 为二值化阈值的迭代次数, m 、 n 为图像的尺寸;基于深度信息的目标区域

细化步骤的时间复杂度为 $O(2mn)$;局部特征提取及多模板匹配步骤的时间复杂度为 $O(T_s mn + 512n_1)$, T_s 与构建差分金字塔步骤的参数有关,在本文算法中 $T_s = 6$, n_1 为特征提取中特征点与伪特征点的总和;均值漂移聚类分割步骤的时间复杂度为 $O(n_2^2)$, n_2 为与模板匹配后保留的有效特征点。本文算法的总时间复杂度为 $O[(T_b + 2 + T_s)mn + 512n_1 + n_2^2]$, 由于 $512 > n_1 \gg n_2$, 故可简化为 $O(mn)$ 。

文献[5]算法为基于颜色的均值漂移过程,其时间复杂度为 $O(T_2 mn)$, T_2 与图像中颜色的种类有关,可简化为 $O(mn)$ 。

FCN 算法^[6]是具有多个卷积层的卷积神经网络,对于每个卷积层 $O(T_i k^2 cm_i n_i)$, m_i 、 n_i 为该层输入图像尺寸, c 为输入图像通道数, k 为卷积核大小, T_i 正比于该层所需要提取的特征的个数,故 FCN 算法的整体时间复杂度为 $O(\sum T_i k^2 cm_i n_i)$ 。每一卷积层的输入图像尺寸 m_i 、 n_i 与算法的原始输入图像尺寸 m 、 n 成正比,故可简化为 $O(mn)$ 。

综上可得,本文算法的时间复杂度与现有算法的时间复杂度相同。

本文算法各阶段的平均运行时间如表 4 所示。含有预处理步骤的算法比不含预处理步骤的算法节省 38.8% 的运行时间。优化后 SIFT 特征提取及匹配步骤仍然为耗时最多的步骤,后续将继续研究如何进一步加速这个步骤。

表 4 本文算法各阶段运行时间(单位:s)

Table 4 Runtime of each stage of the proposed algorithm (unit: s)

Algorithm	Preprocessing	Stereo matching	Region refinement based on depth information	SIFT matching	Mean shift	Total
With preprocessing	0.38	0.06	0.06	0.49	0.02	1.01
Without preprocessing		0.06	0.06	1.51	0.02	1.65

4 结 论

研究了应用于视觉机器人的图像分割算法。针对复杂颜色场景存在的过分割现象,提出了一种基于局部特征的多目标分割算法。在搭建的实验环境中,将本文算法与 FCN 算法、文献[5]算法进行了比较。实验结果表明,在包含复杂颜色信息的多目标场景中,文献[5]算法不适用于搭建的实验场景,而本文算法可以得到与 FCN 算法相近的分割结果。与 FCN 算法相比,本文算法无需大量数据进行预先训练,仅需要少量目标模板就可以在较短的时间内

获得令人满意的分割结果。在视觉机器人的实际使用中对实时性要求较高,如何进一步加速将作为后续研究方向。

参 考 文 献

- [1] Bai L B. Research on vision guidance method of industrial robot[D]. Shenzhen: Shenzhen University, 2016. 白龙彪. 工业机器人的视觉引导方法研究[D]. 深圳: 深圳大学, 2016.
- [2] Wu Q Q. Application of mean shift algorithm in color image filtering and segmentation[D]. Xi'an: Shaanxi

- Normal University, 2015.
- 吴琴琴. Mean Shift 算法在彩色图像滤波与分割中的应用[D]. 西安: 陕西师范大学, 2015.
- [3] Feng J M, Lu Z M, Yang P, *et al.* Color image segmentation algorithm based on local priority [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2014, 42(9): 44-48.
- 冯进玫, 卢志茂, 杨朋, 等. 基于局部占优度的彩色图像分割算法 [J]. 华中科技大学学报(自然科学版), 2014, 42(9): 44-48.
- [4] Lie J, Wang L H, He Y Q, *et al.* Image segmentation method for robot vision [J]. System Engineering and Electronics, 2017, 39(7): 1653-1659.
- 雷俊, 王立辉, 何芸倩, 等. 适用于机器人视觉的图像分割方法 [J]. 系统工程与电子技术, 2017, 39(7): 1653-1659.
- [5] Wang Z Y, Li G, Hou Z Q, *et al.* Color image segmentation by fusion of mean shift and region saliency [J]. Computer Engineering and Applications, 2010, 46(28): 181-184.
- 王朝英, 李光, 侯志强, 等. 融合 mean shift 和区域显著性的彩色图像分割算法 [J]. 计算机工程与应用, 2010, 46(28): 181-184.
- [6] Shelhamer E, Long J, Darrell T. Fully convolutional networks of semantic segmentation [J]. IEEE Transactions on Pattern Recognition and Machine Intelligence, 2017, 39(4): 640-651.
- [7] Feng J W, Zhang L M, Deng X Y. Image segmentation based on multi-source fusion FCN [J/OL]. (2017-10-10) [2017-10-20]. <http://kns.cnki.net/kcms/detail/51.1196.TP.20171010.1732.110.html>.
- 冯家文, 张立民, 邓向阳. 基于多源融合 FCN 的图像分割 [J/OL]. (2017-10-10) [2017-10-20]. <http://kns.cnki.net/kcms/detail/51.1196.TP.20171010.1732.110.html>.
- [8] Zhang Q, Lu S Q, Li H B, *et al.* Research on underwater stereo matching method based on color segmentation [J]. Acta Optica Sinica, 2016, 36(8): 0815001.
- 张强, 卢士强, 李海滨, 等. 基于色彩分割的水下立体匹配算法的研究 [J]. 光学学报, 2016, 36(8): 0815001.
- [9] Huo X Y, Liu Y. A stereovision measurement method using epipolar constraint to correct digital image correlation matching [J]. Chinese Journal of Lasers, 2017, 44(8): 0804003.
- 霍晓洋, 刘洋. 一种极线约束修正数字图像相关匹配的立体视觉测量方法 [J]. 中国激光, 2017, 44(8): 0804003.
- [10] Bai T Z, Hou X B. An improved image matching algorithm based on SIFT [J]. Transactions of Beijing Institute of Technology, 2013, 33(6): 622-627.
- 白廷柱, 侯喜报. 基于 SIFT 算子的图像匹配算法研究 [J]. 北京理工大学学报, 2013, 33(6): 622-627.
- [11] Yang S, Xia M H, Zheng Z S. Medical image registration algorithm based on polynomial deterministic matrix and SIFT transform [J]. Laser & Optoelectronics Progress, 2016, 53(8): 081002.
- 杨飒, 夏明华, 郑志硕. 基于多项式确定性矩阵的 SIFT 医学图像配准算法 [J]. 激光与光电子学进展, 2016, 53(8): 081002.
- [12] Yu Z J, Wang S B. Improved PCA-SIFT algorithm for matching stereo system [J]. Laser & Optoelectronics Progress, 2016, 53(3): 031501.
- 于之靖, 王韶彬. 改进 PCA-SIFT 算法的立体匹配系统 [J]. 激光与光电子学进展, 2016, 53(3): 031501.
- [13] Zhang X, Jin Y X, Xue D. Image matching algorithm based on SICA-SIFT and particle swarm optimization [J]. Laser & Optoelectronics Progress, 2017, 54(9): 091002.
- 张鑫, 靳雁霞, 薛丹. SICA-SIFT 和粒子群优化的图像匹配算法 [J]. 激光与光电子学进展, 2017, 54(9): 091002.
- [14] Pan C, Cui F. Color image segmentation based on spatial and temporal difference sampling [J]. Computer Engineering, 2012, 38(13): 199-201.
- 潘晨, 崔凤. 基于空间和时间差别采样的彩色图像分割 [J]. 计算机工程, 2012, 38(13): 199-201.