

基于视觉统计概率模型的目标定位

谢 昭 高 隽

(合肥工业大学计算机与信息学院图像信息处理研究室, 合肥 230009)
(中国科学院合肥智能机械研究所仿生感知与控制研究中心, 合肥 230031)

摘 要 复杂场景中的目标定位是目标检测和识别的重要过程,为了更好地对复杂场景中的目标进行定位,基于视觉的概率模型,提出了一种目标定位的新方法。区别于一般的区域分割和边缘检测方法,该方法首先通过建立平滑、纹理、阴影和杂乱等4种不同类型区域特性的概率模型,对场景中的前景和背景进行了概率分析;然后结合不同的尺度大小,标记出图像中显著度较高的目标区域;最后经过边缘轮廓的概率建模和连通性分析来提取完整目标区域。实验结果表明,该方法具有较好的鲁棒性和通用性,不仅符合人的视觉注意特性,而且具有一定的抗背景干扰能力。

关键词 视觉概率模型 区域分割 边缘检测 尺度

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2007)07-1234-09

Object Localization Based on Visual Statistical Probabilistic Models

XIE Zhao, GAO Jun

(Department of Computer and Information, Laboratory of Image Information Processing Hefei University of Technology, Hefei 230009)
(Centre for Biomimetic Sensing and Control Research, Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031)

Abstract Object localization in complex settings is the main process in vision task of detection and recognition. This paper presents a new method for object localization based on visual statistical probabilistic models. It is different from traditional ways of region segmentation and edge detection. First, the method can label the regions with high saliency through regional probabilistic models which were developed using the flat, texture, shading and clutter properties with different scales and then the whole object regions can be extracted by edge probabilistic models and connectivity discussion. Experiments show the approach has strong robustness and generality, because the results are in accord with visual attention properties and the "background" noise is lower.

Keywords visual probabilistic models, region segmentation, edge detection, scale

1 引 言

复杂场景中的目标检测和识别是视觉处理的核心问题,也是图像理解的重要环节。目标定位旨在滤去图像中背景的干扰,分割出场景中的物体,并确定其在场景中的位置。特定目标(人脸,汽车等)的

检测^[1-3]是先利用其特殊的结构特征进行样本训练和二值分类,然后用规则形状(通常是矩形或圆形)框定出场景中的目标区域,这不但训练耗时,且检测区域较难反映出目标的真实边界。基于 bottom-up 的分割方法^[4,5]由于受到环境“噪声”的影响,难以有效地标记分割出目标边界,而基于 top-down 的分割方法^[6]则基本上是采用监督策略通过提取训练

基金项目:国家自然科学基金项目(60375011, 60575028);安徽省优秀青年科技基金项目(04042044);“新世纪优秀人才支持计划资助”项目(NCET-04-0560)

收稿日期:2006-01-26;改回日期:2006-04-16

第一作者简介:谢昭(1980~),男。现为合肥工业大学计算机与信息学院在读博士研究生。主要研究方向为图像理解、智能计算方法等。E-mail: xiezhaol980@126.com

样本的局部信息来指导分割,但其对每类样本的获取则比较困难,因此缺乏先验信息时,如何在复杂场景中找出完整的目标边界非常重要。

视觉选择性注意机制的研究是通过对人眼的视觉特性进行分析^[7,8],并结合图像的颜色、方向和纹理信息进行显著图的计算,标记出复杂场景中的显著区域,其不仅构建的计算函数比较复杂,且得到的结果基本上是 bottom-up 的数据驱动,并缺乏知识信息的指导。本文采用注意机制的基本思想,先利用视觉统计方法建立了 4 类特征区域的概率模型,然后结合图像分割和边缘提取的方法进行了尺度分析,并形成了有效的指导策略,从而构建了完整的粗定位区域分割和细定位边缘提取框架。实验是采用若干自然场景灰度图进行分析和评价,结果表明,目标定位的效果基本上符合视觉感知,而且保留了较完整的边缘信息和形状特征。

2 基于视觉统计模型的目标定位框架

本文提出的目标定位方法,在图像空间中采用由粗至精的搜索策略 (coarse-to-fine searching),基本的目标定位框图如图 1 所示,其中包括粗略定位和精确定位两部分。原始图像经过预处理后,再进行多尺度的区域局部统计。目标分割和定位时,先计算区域的平滑特性、杂乱特性、纹理特性及其阴影特性的概率值,然后经过概率分析组合,只选择符合人眼视觉感知的高概率尺度区域,以实现目标的粗略定位;接着再根据不同的边缘检测算子对粗定位区域进行边缘提取,同时通过边缘概率分析滤去假边缘干扰;最后利用边界连接策略形成目标轮廓的闭合曲线,即完成目标的定位和分割。

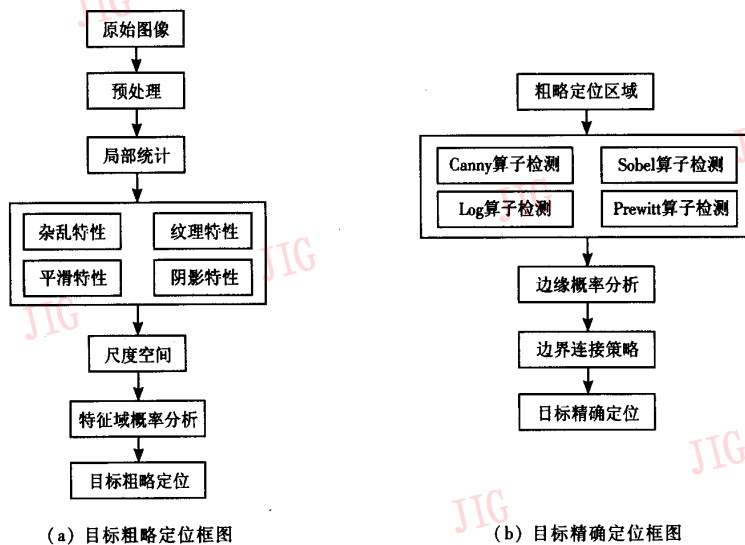


图 1 目标定位框架
Fig. 1 The frames of object localization

3 区域统计概率模型的粗略定位

为了能精确地标记出复杂背景中的目标轮廓,需要对整幅图像场景中的目标进行粗略定位,以便减少目标定位过程中的搜索空间。

人之所以能较快地感知复杂环境中的目标,是因为视觉具有相对性和选择性特征^[9],由于对前景和背景的区别往往取决于前景与背景的对比值,而

前景区域能给人更为明显的视觉刺激,具有更为显著的灰度、纹理和曲线特征,因此局部一致性更为明显;人视觉的选择性特性则能有效地过滤出背景的噪声干扰。本节利用人的视觉特性通过构建区域概率模型来进行目标区域的概率分析和粗略定位。

3.1 区域类型统计概率模型

设 W, H 分别为图像的宽度和高度,则 2 维图像空间定义为

$$\Omega = \{(i, j) : 1 \leq i \leq W, 1 \leq j \leq H\} \quad (1)$$

对空间中的任意点 $\nu \in \Omega$, 其灰度值 $I_\nu \in \{0, \dots, G\}$, G 取 255, 粗定位就是在若干个区域 K_i 中确定目标所在的大致区域范围, 文献[10]详细阐述了基于区域块的中层复杂度特征能最有效地进行目标的分割和识别, 因此对场景中的区域块进行概率分析非常有意义。

通过对大量图像的区域研究表明, 有 4 类区域在真实世界的图像中出现的频率较高^[11], 如图 2 所示的 4 幅图像中的窗口区域(白线框区), 其中图 2(a)为图像中的平滑区域, 图 2(b)为图像中的杂乱区域, 图 2(c)显示了图像中的纹理区域, 图 2(d)则显示了图像中的阴影区域。

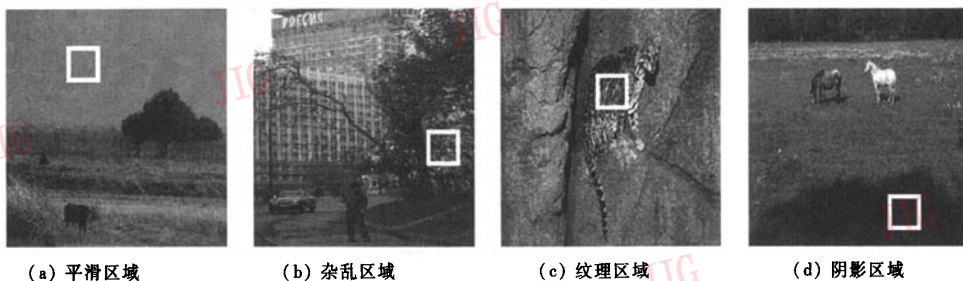


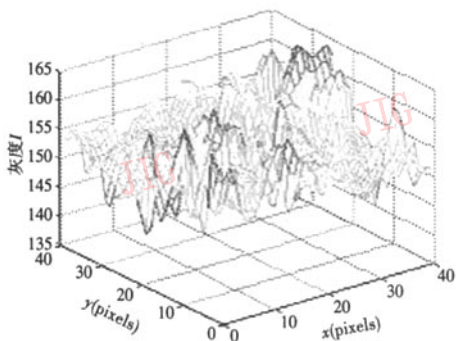
图 2 4 种不同类型的窗口区域

Fig. 2 Different four types of regions

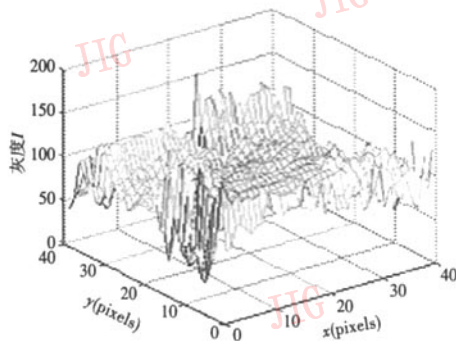
与 4 种特征区域对应的灰度空间分布如图 3 所示, 由图 3 可见, 平滑区域 r_1 灰度变化幅度小, 其灰度值服从高斯分布, 区域 K 为平滑区域的概率为

$$p_1(K) = \prod_{\nu \in K} G(I_\nu - \mu; \sigma^2) \quad (2)$$

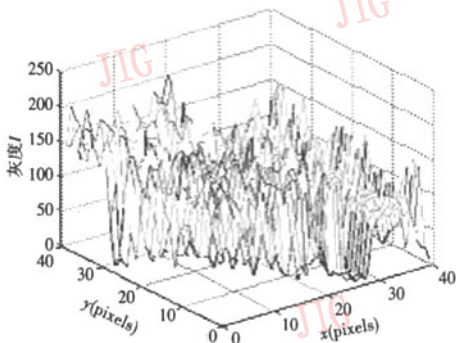
式中, 参数 μ, σ 分别为区域 K 的灰度均值和方差。



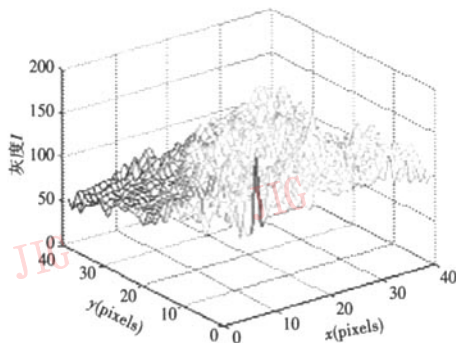
(a) 平滑区域空间灰度图



(b) 杂乱区域空间灰度图



(c) 纹理区域空间灰度图



(d) 阴影区域空间灰度图

图 3 4 种类型区域的空间强度图

Fig. 3 Spatial intensity map of four type regions

因图像杂乱区域 r_2 中的灰度没有明显分布特征,且灰度分布幅值很大,呈现随机特性,故对该区域的概率建模可采用如式(3)所示的区域局部直方图进行统计:

$$p_2(K) = \prod_{v \in K} h(I_v) = \prod_{j=0}^G h_j \quad (3)$$

其中, h_j 为区域中的灰度值为 j 的概率值,其可由区域中灰度值为 j 的像素个数除以区域总个数得到。

图像纹理区域 r_3 的灰度空间分布呈现出明显的重复性和方向性特征,因此 Gabor 滤波器^[12]对像素邻域的灰度属性进行统计分析能有效地得到纹理的方向性和粗糙度。本文采用 8 个方向的变尺度 Gabor 滤波来对纹理区域进行统计分析。令 $v \in K$ 为滤波器的中心, Φ_v 为点 v 的尺度邻域, $H(I_v)$ 为中心在点 v 处的 8 组滤波器局部直方图的统计概率, $H(I_{\Phi_v})$ 为中心在 Φ_v 范围,其窗口覆盖点 v 处的所有区域滤波器局部直方图的统计概率,则区域 K 为纹理区域的概率如下:

$$p_3(K) = \prod_{v \in K} \frac{1}{\nu_{\text{normal}}} \exp\{-\langle \Theta, H(I_v | I_{\Phi_v}) \rangle\} \quad (4)$$

Θ 为区域局部直方图统计参数, $H(I_v | I_{\Phi_v}) = H(I_v) / \sum_{v \in \Phi_v} H(I_v)$, $\langle \cdot, \cdot \rangle$ 为向量的内积, ν_{normal} 是向量 ν 归一化处理后的结果。

由于图像阴影区域 r_4 中的灰度特征符合非一致性区域的灰度渐变特性,因此可利用 2 维 Bézier 曲面进行建模,对于区域中任意的点 $v = (x, y) \in K$, 令 $B(x, y)$ 为 Bézier 曲面,则

$$B(x, y) = P_{(x)}^T \times M \times P_{(y)} \quad (5)$$

其中, $x, y \in (0, 1)$, $P_{(x)} = ((1-x)^3, 3x(1-x)^2, 3x^2(1-x), x^3)^T$ 为 Bézier 基函数, $P_{(y)}$ 类似, M 是与由 16 个固定的区域控制点组成的 4×4 维的空间点列对应的灰度矩阵,其点列由最靠近中心线的两条纵横直线与原窗口交汇的 16 个交点构成(如图 4 所示),其中点 O 为窗口中心点,其余点为选取的控制点空间点列,区域 K 的阴影区域的概率服从高斯分布(如式(6)所示),其中 σ 为区域的灰度方差, B_v 为与区域点对应的 Bézier 曲面值。

$$p_4(K) = \prod_{v \in K} G(I_v - B_v; \sigma^2) \quad (6)$$

3.2 尺度空间下的类型区域概率分析

图像中的目标区域具有随机性和不确定性,而目标本身的差异性则导致了粗定位的过程需要在尺度空间下通过综合不同类型的特征区域概率模型来

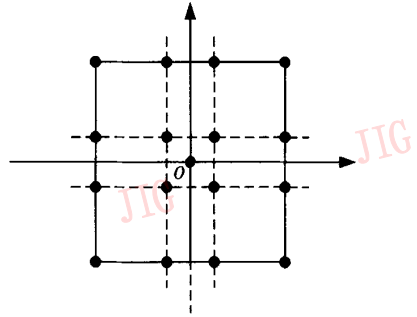
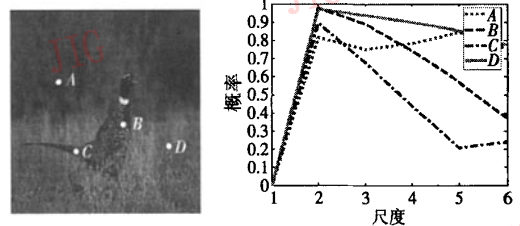


图 4 Bézier 曲面控制点选取

Fig. 4 The selection of control points in Bézier surface

进行区域分析和取舍。对图像中目标定位的区域块选取,本文采用的变尺度窗口的大小为 $(4n+1) \times (4n+1)$,其中 $1 < n \leq 6$,一般的,由于大尺度区域统计保证了图像中目标的完整性,因此不会造成目标的遗漏;而小尺度区域统计则不仅保证了图像中目标的精确性,而且降低了目标以外的噪声干扰。

高概率的平滑区域可能包括大范围的背景区域和小尺度范围内的目标区域。窗口中心位于 A, B, C, D 点, $n=2$ 的小尺度高概率平滑区域如图 5(a)所示,图 5(b)所示为 4 点中心区域随着尺度的增加,其平滑概率的变化情况,由图 5(b)可见,以 A, D 为中心点的大面积背景平滑区域,其概率特性随着尺度的增加并未有明显的衰减,而以 B, C 为中心点的小面积目标区域则由于受到目标边缘的影响,其概率呈明显下降趋势,因此在平滑区域中可以根据其概率的下降程度将大面积的背景平滑区域去除。



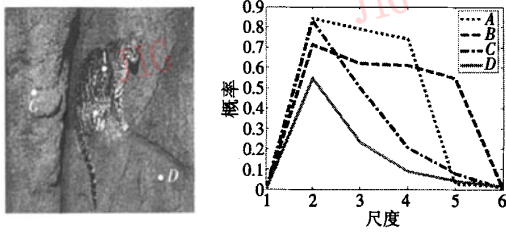
(a) 包含 4 点区域的原图像

(b) 尺度空间中平滑区域概率变化曲线

图 5 尺度空间中平滑特性的概率分析

Fig. 5 Probability analysis for flat property in scale space

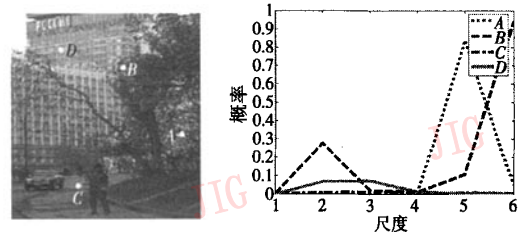
高概率的纹理区域能反映出纹理特性丰富的目标区域的位置信息,窗口中心位于 A, B, C, D 点, $n=2$ 的小尺度高概率纹理区域如图 6(a)所示,从图 6(b)中可以看出,在 $2 \leq n \leq 4$ 的尺度范围内,其



(a) 包含 4 点区域的原图像 (b) 尺度空间中纹理区域概率变化曲线

图 6 尺度空间中纹理特性的概率分析

Fig. 6 Probability analysis for texture property in scale space



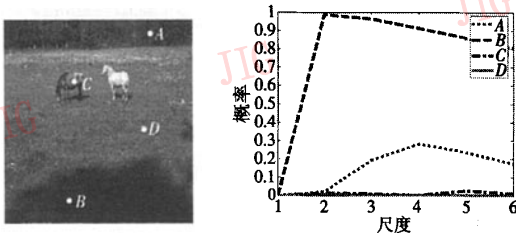
(a) 包含 4 点区域的原图像 (b) 尺度空间中杂乱区域概率变化曲线

图 8 尺度空间中杂乱特性的概率分析

Fig. 8 Probability analysis for clutter property in scale space

目标区域的纹理特性的概率变化幅度很小,而背景区域的纹理特性的概率变化则非常明显,因此在一定尺度范围内可以将概率变化不明显的纹理高概率区域视为具有高纹理特性的目标区域,从背景中提出。

阴影区域的高概率模型能够并且仅能标记因为光照等环境影响而形成的阴影,图 7(b)显示了图 7(a)中随着尺度变化的对应 4 点的阴影区域的概率变化情况,从图 7(b)中可以看出,由于非阴影区域的概率几乎为零,而可能的阴影区域则概率很高,因此对于概率高的阴影区域可以将其视为由目标产生的阴影背景或暗背景。



(a) 包含 4 点区域的原图像 (b) 尺度空间中阴影区域概率变化曲线

图 7 尺度空间中阴影特性的概率分析

Fig. 7 Probability analysis for shading property in scale space

杂乱区域的高概率模型标记了杂乱的背景区域,图 8(b)显示了图 8(a)中随尺度变化的对应 4 点杂乱区域的概率变化情况,从图 8(b)中可以看出,处于杂乱背景区域的中心点总会在某一尺度下表现出极大的杂乱高概率特性,而其他区域的概率则随着尺度变化而变化平缓,因此对杂乱的高概率区域可视为背景。

在通常情况下,场景中的目标主要包括小区域的平滑区域和纹理区域,而阴影区域和杂乱区域则

是背景区域,根据对 4 种类型区域的分析,本文建立了如下的目标区域总概率模型:

$$p(K | \tau_i) = w_1 p_1(K | \tau_i) - w_2 p_2(K | \tau_i) + w_3 p_3(K | \tau_i) - w_4 p_4(K | \tau_i) \quad (7)$$

$\tau_i (i=2,3,4,5,6)$ 为尺度大小, $w_j (j=1,2,3,4)$ 分别为各区域模型的权重值, $w_j = \lambda_j \cdot \frac{\bar{p}_j}{\sum_{i=1}^6 \bar{p}_j}$, λ_j 为尺度

影响因子, \bar{p}_j 为区域内的概率均值,其中

$$\lambda_1 = \sum_{i=2}^6 \frac{1}{1 + e^{-(p_1(K, \tau_i) - p_1(K, \tau_{i-1}))^2}} p_1(K | \tau_i) \quad (8)$$

$$\lambda_2 = \max_{1 \leq i \leq 6} (p_2(K | \tau_i)) \quad (9)$$

$$\lambda_3 = \sum_{i=2}^6 e^{-(p_3(K, \tau_i) - p_3(K, \tau_{i-1}))^2} p_3(K | \tau_i) \quad (10)$$

$$\lambda_4 = \max_{1 \leq i \leq 6} (p_4(K | \tau_i)) \quad (11)$$

$$\bar{p}_j = \frac{1}{M \times N} \sum_{\kappa \in \Omega} p_j(\kappa) \quad (12)$$

M, N 为图像大小;权重中的尺度因子反映出同一区域不同尺度下变化情况,而式(8)和式(10)的 Sigmoid 函数和指数函数则反映了平滑区域和纹理区域相邻尺度间的概率变化对权重的影响,其也降低了大面积背景平滑区域和小面积背景纹理区域所属前景区域的概率;场景的概率均值则反映出整幅场景中 4 种类型的概率值所占的比重,例如,若场景中纹理特性的概率所占比重较少,则场景中包含有纹理的目标区域的概率也较少,从而减小了纹理区域的权重;同样的,若场景中的杂乱区域所占比重较大,则相应区域为目标区域的概率降低,此时可过滤背景区域。

3.3 粗定位算法描述

粗定位的过程就是通过区域前景概率分析来进行目标判断的过程,以便将大面积的背景区域从场

景中过滤。区域统计模型的粗定位算法的详细步骤描述如下:

for $i = 1$ to M

for $j = 1$ to N

for $n = 2$ to 6

- (1) 确定区域 K , 其窗口中心位于 (i, j) , 尺度大小为 $(4n + 1) \times (4n + 1)$;
- (2) 由式(2) ~ 式(6)分别计算该区域的 4 种类型区域的概率 $p_k(K | \tau_n), k = 1, 2, 3, 4$;
- (3) 由式(8) ~ 式(12)计算各类型区域的概率权重值 $w_k, k = 1, 2, 3, 4$;
- (4) 由式(7)计算目标区域的总概率值 $p(K | \tau_n)$;
- (5) 确定候选的目标区域 Ω_{k_0} 满足

$$\Omega_{k_0} = \bigcup_{K \subset \Omega} \bigcup_{n \in \{2, 3, 4, 5, 6\}} p(K, \tau_n) \geq T$$

T 为基本概率阈值, 其选取采用空间聚类方法, 将概率空间分为前景和背景两类, 令 ν_B 为背景概率聚类中心值, ν_0 为目标概率聚类中心值, 则基本概率阈值选取为 $(\nu_B + \nu_0)/2$ 。

end

end

end

令 $\Omega_{r_1}, \Omega_{r_2}, \Omega_{r_3}, \Omega_{r_4}$ 为 4 种高概率特征区域集

合空间, Ω_k 为最终的粗定位目标区域, $\Omega_k \subset \Omega_{k_0}$, 采用本文提出的粗定位算法对目标区域进行判断, 遵循了以下几种情况:

- (1) 杂乱高概率区域为背景区域, 即 $\Omega_{r_2} \subset \Omega - \Omega_k$;
- (2) 阴影高概率区域为背景区域, 即 $\Omega_{r_4} \subset \Omega - \Omega_k$;
- (3) 既是平滑高概率区域, 又是纹理高概率的区域为粗定位的目标核心区域, 即 $\Omega_{r_1} \cap \Omega_{r_3} \subset \Omega_k$;
- (4) 单一的平滑高概率区域或纹理高概率区域与粗定位的核心区域邻接, 且具有与核心区域的高概率类型类似的特征区域, 则是目标区域;
- (5) 其他情况下, 大面积平滑区域和小范围纹理区域为背景区域。

图 9 显示了经过粗定位后的部分目标区域定位结果, 其中背景区域均变为白色。从图 9 可以看出, 粗定位过程基本上保留了图像中显著的目标区域, 其不仅有效过滤了背景的干扰, 而且减小了精确定位的搜索空间。

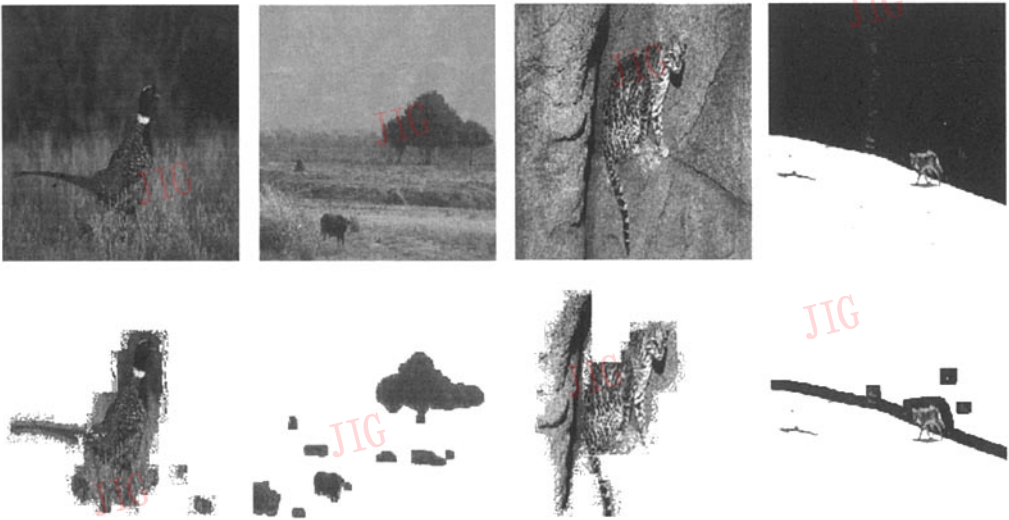


图 9 原始图和粗定位过程的结果

Fig. 9 Original images and results in rough localization process

4 边缘概率模型的精确定位

经过区域概率模型初步定位的目标不仅符合视觉的“前景”特性, 即具有高显著性和高分度, 而且处于概率极值区域, 这就保证了区域目标的完整。但由于整个过程为盲目搜索, 若采用规则窗口, 则不能

精确地区分和表达客观世界中真实目标千差万别的复杂轮廓形状, 因此, 需要在已粗定位划分的目标区域进一步区分“前景”和“背景”的轮廓; 为此, 本文通过建立边缘概率模型来对目标进行边缘提取和精确定位。

令 Ω_k 为粗定位区域空间, $\Psi^{(n)}$ 为边缘检测算子空间, n 为算子个数, 考虑到算法的通用性, 这里采用了 Canny 算子、Sober 算子、Prewitt 算子和 Log

算子 4 种常用的边缘检测算子进行边缘检测, $\Omega_{\partial K}$ 为目标轮廓空间, $\Omega_{\partial K}$ 可分为以下 3 类:

(1) 假边缘空间 ∂r_1 , 即由粗定位引入的原图像中并不真实存在的边缘空间, 这种边缘不是目标的真实轮廓;

(2) 弱边缘空间 ∂r_2 , 即由粗定位未正确划分的“背景”对目标造成的边缘影响, 或目标本身的不显著轮廓, 这种边缘可能是目标的真实轮廓;

(3) 强边缘空间 ∂r_3 , 即目标区域中区域区分度较为明显的轮廓, 这种边缘既可能是目标的真实轮廓, 又可能是目标内部不同区域的轮廓。

令 ∂K 为边缘算子检测到的曲线段或直线段, $\partial K \subseteq \Omega_{\partial K}$, 则曲线段属于假边缘的概率如下:

$$p(\partial r_1 | \partial K) = \begin{cases} 0; & \text{if } \partial K \subseteq \Omega_K \cap \Omega_{\partial K} \\ 1; & \text{if } \partial K \not\subseteq \Omega_K \cap \Omega_{\partial K} \end{cases} \quad (13)$$

如果舍弃假边缘线段, 并进一步排除了背景噪声的干扰, 那么就可弱边缘和强边缘空间中, 将边缘点集合组成的空间 V 定义为 $V = \{\nu | \nu \in \Omega \cap \nu \in \Omega_{\partial K}\}$, 即

由各种边缘算子检测出的边缘点集合, 对任意的点 $\nu \in V$, 空间中的边缘点归属于目标轮廓的概率为

$$p(\nu | \Omega_{\partial K}) = \prod_{\varphi \in \Psi^{\nu}, \nu \in \partial K^{\nu}} \frac{l^{\nu}(\partial K^{\nu})}{\max(l^{\nu}(\partial K^{\nu}))} \quad (14)$$

φ 为边缘检测算子子空间集合, $l^{\nu}(\partial K^{\nu})$ 表示边缘算子子空间中 ν 点所在边缘线段的八邻域链码长度, $\max(l^{\nu}(\partial K^{\nu}))$ 表示边缘算子子空间中的最长链码。

然后就可对边缘轮廓空间点的概率进行聚类, 将概率空间分为强边缘点和弱边缘点两类, 令 ν_w 为弱边缘概率聚类中心值, ν_s 为强边缘概率聚类中心值, 若基本概率阈值选取 $(\nu_w + \nu_s)/2$, 则强边缘点的概率大于阈值, 而弱边缘点的概率则小于阈值。强边缘点属于目标的真实边缘点, 而弱边缘点则作为进一步滤去的边缘点滤除。图 10 显示了某目标的强边缘、弱边缘和假边缘点集合。边界连接策略用于对强边缘中的短线段进行连接, 通过直线连接距离主轮廓边缘最近的端点(八连通)形成最终的闭合曲线, 即可实现场景中目标和背景的分。

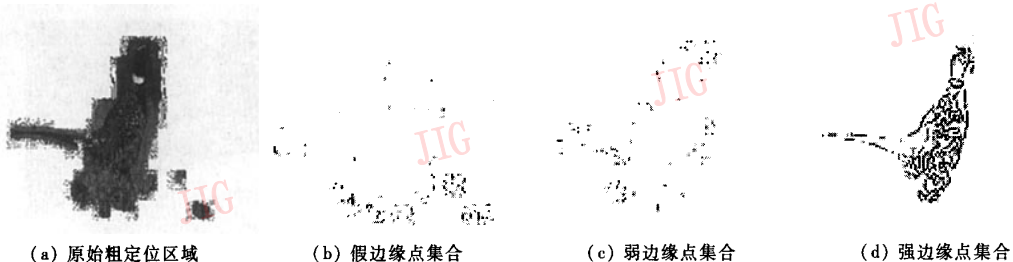


图 10 不同类型的边缘概率点集合

Fig. 10 Different types of point sets with edge probability

5 实验结果

为验证本文算法的定位分割效果, 利用 Berkeley 图像库^[13]中的若干场景灰度图像进行了相关实验, 图像大小为 256×256 。图 11 显示了原图和最终分割结果, 从图中可以看出, 最终的分割结果不仅有效地排除了背景的干扰, 得到了目标的真正的轮廓, 而且符合人眼视觉的显著性特点。

图 12 是利用经典的 Neuts 分割算法^[5]对若干幅不同图像进行分割的结果。

从图 11 和图 12 中可以看出, 本文提出的分割定位算法能更为有效地实现前景和背景的分, 不仅可保证前景目标的完整性, 而且可减小背景噪声的干

扰。表 1 列出了两种算法得到的与前景区域对应的 4 种不同特征模型在不同尺度下的概率极大值, 其中 Neuts 算法的前景区域为与真实区域重叠面积最大的分割区域, 括号中数据对应概率极大值时的尺度 σ_i 。由于 Neuts 算法不能较完整地分离出目标区域, 因此各种类型区域的概率比较平均, 而新方法的概率极大值则比较集中在平滑区域或纹理区域中。

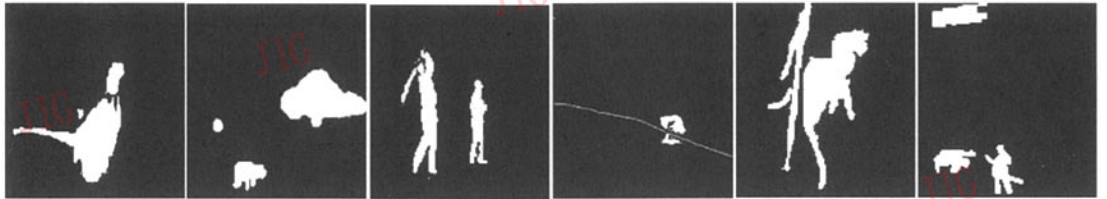
表 2 是两种算法对目标分割的定量评价, 其中, 准确度 A 和完整度 U 标准分别描述了目标定位过程中的抗背景“噪声”特性和前景“一致”特性, 其定义如下:

$$A = \frac{S}{R} \quad U = \frac{R}{S} \quad (15)$$

R 表示定位分割后包含目标的区域面积, S 表示图



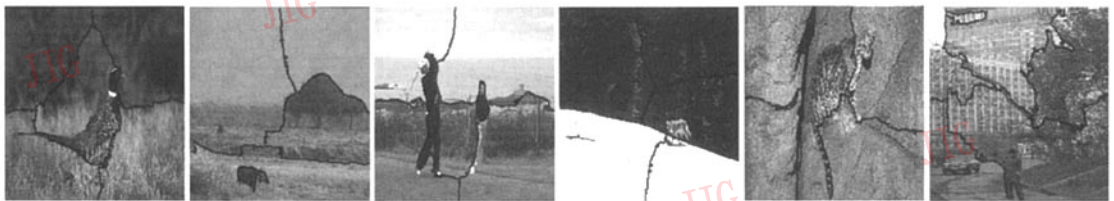
(a) 原图像



(b) 分割结果

图 11 基于视觉概率模型的目标定位实验

Fig. 11 Experiments of object localization based on visual statistical probabilistic models



(a)

(b)

(c)

(d)

(e)

(f)

图 12 Ncuts 分割算法的分割结果

Fig. 12 Segmentation results using Ncuts algorithm

表 1 新算法与 Ncuts 算法分割的目标区域在不同尺度下对应 4 种特性的概率均值的极大值

Tab. 1 Highest average probability in different types of regions corresponding scales with proposed method and Ncuts

场景图像	Ncuts 算法				新算法			
	平滑	杂乱	纹理	阴影	平滑	杂乱	纹理	阴影
BIRD	0.26(3)	0.32(3)	0.37(3)	0.28(5)	0.59(3)	0.22(4)	0.31(5)	0.18(4)
BULL	0.45(4)	0.49(4)	0.34(5)	0.24(3)	0.57(2)	0.35(4)	0.23(3)	0.09(5)
CAT	0.41(3)	0.23(5)	0.57(4)	0.27(3)	0.65(2)	0.23(3)	0.75(5)	0.19(4)
GOLF	0.35(3)	0.40(3)	0.27(3)	0.34(3)	0.51(4)	0.12(3)	0.07(4)	0.34(4)
SCENE	0.17(2)	0.08(3)	0.14(3)	0.25(5)	0.54(3)	0.24(3)	0.10(3)	0.29(4)
WOLF	0.51(2)	0.14(3)	0.32(4)	0.66(3)	0.53(3)	0.18(4)	0.07(5)	0.44(3)

表 2 新算法与 Ncuts 算法的目标分割定位结果比较

Tab. 2 Result comparison between proposed method and Ncuts in object location and segmentation

场景图像	目标个数	Ncuts 算法		新算法	
		A	U	A	U
BIRD	1	0.162	-	0.957	-
BULL	3	0.474	0.976	0.943	-
CAT	2	0.334	-	0.988	0.845
GOLF	2	0.445	-	0.981	-
SCENE	3	0.143	-	0.764	0.904
WOLF	1	0.108	-	0.892	-

像库中手工标记图像的区域面积,当分割区域包含背景区域时,用 A 衡量;当分割区域仅包括目标区域时,用 U 衡量,当 $A \rightarrow 1, U \rightarrow 1$ 时,则能最准确描述真实目标。

从表 2 中可以看出,新算法较 Ncuts 算法能有效地标记出完整的目标区域,其受背景噪声干扰小,且准确度 A 的数值是两者之间的主要误差,Ncuts 分割过程中将目标区域分割为背景区域和将背景区域同时分割为目标区域的情况较为常见,新方法能有

效避免这种情况的发生。

6 结论

本文结合视觉统计模型提出了一种用于目标分割的新方法,该算法通过建立区域和边缘概率模型,能有效标记出复杂场景中显著度较高的目标区域和轮廓,不仅抗背景干扰能力强,且符合人的视觉注意特性,这就为图像理解的识别过程提供了完整的目标对象和分析样本,另外还通过相关对比实验证明了方法的可操作性和通用性。

参考文献 (References)

- 1 Viola P, Jones M. Rapid object detection using a boosted cascade of simple features [A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C], Hawaii, USA, 2001: 511 ~ 518.
- 2 Wang Hai-chuan, Zhang Li-ming. A novel fast training algorithm for Adaboost [J]. Journal of Fudan University (Natural Science), 2004, 43(1): 27 ~ 33. [王海川, 张立明. 一种新的 Adaboost 快速训练算法 [J]. 复旦学报 (自然科学版), 2004, 43(1): 27 ~ 33.]
- 3 Schneiderman H, Kanade T. Object detection using the statistics of parts [J]. International Journal of Computer Vision, 2004, 56(3): 151 ~ 177.
- 4 Ye Qi-xiang, Gao Wen, Wang Wei-qiang, et al. A colour image segmentation algorithm using colour and spatial information [J]. Journal of Software, 2004, 15(4): 522 ~ 530. [叶奇祥, 高文, 王伟强等. 一种基于图像颜色和空间信息的彩色图像分割算法 [J]. 软件学报, 2004, 15(4): 522 ~ 530.]
- 5 Shi J, Malik J. Normalized cuts and image segmentation [A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C], San Juan, Puerto Rico, 1997: 731 ~ 737.
- 6 Borenstein E, Ullman S. Class specific top down-segmentation [A]. In: Proceedings of the European Conference on Computer Vision [C], Copenhagen, Denmark, 2001: 110 ~ 122.
- 7 Sun Y R, Fisher R. Object-based visual attention for computer vision [J]. Journal of Artificial Intelligence, 2003, 146(1): 77 ~ 123.
- 8 Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254 ~ 1259.
- 9 Gao Wen, Chen Xi-ling. Computer vision [M]. Beijing: Tsinghua University Publisher Press, 1999. [高文, 陈熙霖著. 计算机视觉 [M]. 北京: 清华大学出版社, 1999.]
- 10 Ullman S, Vidal-Naquet M, Sali E. Visual features of intermediate complexity and their use in classification [J]. Nature Neuroscience, 2002, 5(7): 682 ~ 687.
- 11 Zhu S C. Statistical modeling and conceptualization of visual patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(6): 1 ~ 22.
- 12 Tu Z W, Zhu S C. Image segmentation by data-driven Markov Chain Monte Carlo [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 657 ~ 673.
- 13 Berkeley image base [EB/OL]. <http://www.cs.berkeley.edu/projects/vision/grouping/segbench/>