

# 基于轮廓分层描述的目标识别算法研究

黄伟国<sup>1</sup>, 顾超<sup>1</sup>, 尚丽<sup>2</sup>, 杨剑宇<sup>1</sup>, 朱忠奎<sup>1</sup>

(1. 苏州大学城市轨道交通学院, 江苏苏州 215131; 2. 苏州职业大学电子信息工程系, 江苏苏州 215000)

**摘要:** 有效的目标轮廓分段是描述目标局部特征的关键环节. 针对现有轮廓描述算法存在轮廓分段不合理的问题, 本文基于认知心理学, 提出了分层描述的轮廓描述算法. 算法思想是首先根据角点特征将整个轮廓划分成一些轮廓分段, 接着对轮廓分段的分布特点提出价值尺度, 然后将多级轮廓分段按照价值尺度原则合并得到有限个能够完整描述目标轮廓的特征分段, 最后将特征分段综合考虑长度尺度应用到 Shape Context 相似度检测模型中进行目标识别. 通过对 MPEG-7 图像数据库中的图像进行实验分析表明该算法能够完整描述目标图像的形状特征, 提高了目标识别率和形状检索率, 并对部分遮挡的目标也具有良好的鲁棒性. 基本满足目标识别和形状检索对准确率、稳定性、抗遮挡能力等方面的要求.

**关键词:** 轮廓描述; 目标识别; 价值尺度; 部分遮挡

**中图分类号:** TP391.4      **文献标识码:** A      **文章编号:** 0372-2112 (2015)05-0854-08

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2015.05.004

## Hierarchical Representation Method for Object Recognition

HUANG Wei-guo<sup>1</sup>, GU Chao<sup>1</sup>, SHANG Li<sup>2</sup>, YANG Jian-yu<sup>1</sup>, ZHU Zhong-kui<sup>1</sup>

(1. School of Urban Rail Transportation, Soochow University, Suzhou, Jiangsu 215131, China;

2. School of Electronics and Information, Suzhou Vocational University, Suzhou, Jiangsu 215000 China)

**Abstract:** Efficient object contour segment is a critical problem to describe the local features of objects. In order to solve the improper contour segments obtained by the existing recognition methods, a hierarchical description algorithm of contour description is proposed. Firstly, the whole contour is divided into several contour segments by the corners on the contour. Then the valuation scale is put forward via the distribution of contour segments. Thirdly, combine these contour segments into several contour feature segments according to the valuation scale. Finally, the similarity of different contour feature segments, in combination with their lengths, is jointly used to get the best recognized results. The experimental results of MPEG-7 database indicate that this algorithm has great advantage over recently published algorithms, especially for the objects with partial occlusion. Hence, this novel algorithm satisfies the requirements of accuracy, robust and anti-occlusion in object recognition and shape retrieval.

**Key words:** contour description; object recognition; valuation scale; part occlusion

## 1 引言

目标的形状描述是计算机图形学和视觉领域研究的热点和难点. 认知心理学专家认为人们对目标的形状特征比颜色和纹理特征更敏感, 识别的鲁棒性和稳定性更高<sup>[1]</sup>. 在目标识别和形状检索中, 轮廓形状具有非常重要和稳定的视觉特征, 能够反映重要的特征信息, 因此学者们提出了许多基于轮廓形状的目标识别算法<sup>[2~6]</sup>.

理解和描述目标轮廓特征在形状匹配领域中是一项基础工作. 根据目标轮廓形状特征描述方法不同, 可

以将描述方法分为全局描述子、局部描述子<sup>[7]</sup>.

全局特征描述符能够表示目标形状的完整信息, 对目标由噪声引起的局部变形有良好的抗干扰能力, 但是捕捉目标的局部细节特征能力不强, 尤其是对局部遮挡的目标识别能力不高. Belogic 等<sup>[2]</sup>提出的形状上下文 (Shape Context, SC) 描述方法, 这种算法着重关注轮廓上某个点与其它所有点的空间分布关系, 并用数学统计法表示. 这样的表示方法含有丰富的信息, 对目标轮廓的描述能力非常强, 但是该方法抑制噪声能力不强, 对发生大变形的目标形状识别精度不高. Ling 等<sup>[8]</sup>在 SC 基本思想指导下, 用轮廓点之间的内距离代替了欧式距

离,提出了内距离形状上下文 (Inner-Distance Shape Context, IDSC) 方法.该方法对非刚性物体轮廓的形状描述具有良好的表示,但对大变形的目标识别精度不高,且计算复杂度较高. Latecki 等<sup>[9]</sup>使用视觉部分 (Visual Part) 对目标轮廓进行描述,该方法使用离散曲线演化 (Discrete Curve Evolution, DCE) 的方法来对形状进行简化,对小变形的目标识别精度较高,并节约匹配运算时间,但易把特征当成噪声删除,发生误匹配.

为了解决全局特征描述子捕捉局部特征能力弱的缺陷,众多学者提出了局部特征描述符描述目标的轮廓形状.局部特征由于提取了轮廓最重要的分段信息,能够克服目标发生局部变形和遮挡带来的识别困难,但计算复杂度较高,且易受轮廓各参数特征影响. Xiang Bai 等<sup>[10]</sup>采用离散曲线演化的方法对目标形状进行简化并获得目标最重要的轮廓分段,提出了形状相似度 (Shape Similarity) 方法.该方法能够识别发生局部遮挡和变形的目标形状,但同样也具有 DCE 的缺陷. Chun-jing Xu 等<sup>[11]</sup>根据轮廓形状凹凸性拓扑关系提出了轮廓灵活性准 (Contour Flexibility),该方法能够准确获得目标轮廓的全局特征和局部特征,对局部变形和遮挡的目标鲁棒性较强,但计算复杂度较高. Kang B. Sun 等<sup>[12]</sup>根据轮廓长度将轮廓分段,并将轮廓分段采用主成分分析方法提取出最能代表目标轮廓的分段,提出了轮廓分段集准则 (Class Segment Sets).该方法能解决遮挡问题,但不能解决局部变形目标的识别.

本文针对现有目标轮廓特征描述方法存在特征描述不完整的问题,根据人类视觉角度的认知心理学原则,评价了轮廓点和轮廓分段所描述特征的重要程度,提出了一种基于分层描述的目标轮廓特征分段的提取方法.该方法能够提高目标的识别准确率和形状检索率,并能有效识别部分遮挡的目标.

## 2 分层描述算法基本原理

不同的轮廓局部特征对目标识别的准确率是不同的,有效的轮廓分段能够高效完整地描述目标的特征,该分段不仅能够描述目标本身还可以用于区分其他目标;而一些冗余的轮廓分段不仅不能完整描述目标轮廓特征,也易与其它目标发生误匹配.图 1 给出了利用两种不同的轮廓分段识别目标的例子,曲线  $S_1$  和  $S$  是从图像 Heart-3 中提取的,作为分类目标,在以下三幅图中寻找出与每一条曲线最匹配的轮廓分段,第二行是轮廓分段  $S_1$  的匹配结果,第三行是轮廓分段  $S$  的匹配结果,在每幅图中的红色曲线是图像中与轮廓分段最相似的部分.由图 1(b)可以看出,轮廓分段  $S_1$  不能完全描述目标的几何特征,因为三幅图像计算出的 SC 距离基本相似;但是轮廓分段  $S$  能够完整描述目标的几

何特征,可以判断出  $S$  来自图像 Heart-3,因为其对应的 SC 距离值与其它两个相差很大,且是三个中最小的.轮廓分段  $S$  称为轮廓特征分段,本文首要工作就是讨论如何将目标轮廓分割成几段轮廓特征分段.

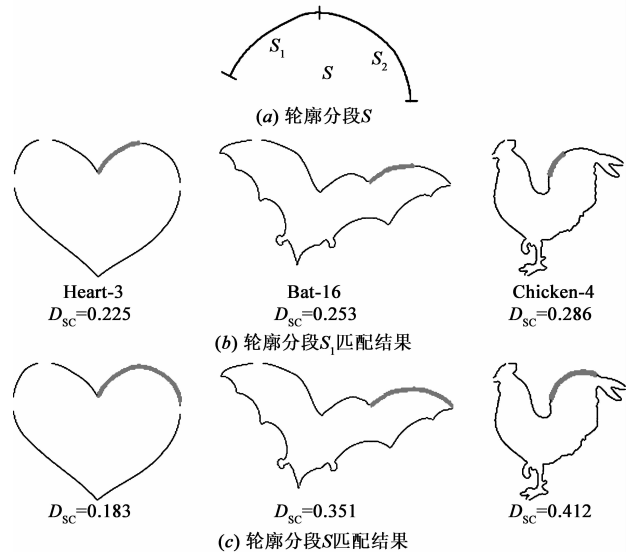


图1 两种不同轮廓分段识别结果

本文从认知心理学视觉角度出发,根据轮廓点曲率分布将轮廓上的每个点分层描述成特征点和非特征点,同时按价值尺度原则将轮廓分段分为轮廓特征分段和非特征分段,并通过合并算法将若干个非特征轮廓合并成凹凸性更明显的特征分段,增强了目标轮廓描述的完整性和高效性,提高目标识别准确率和形状检索率.

### 2.1 轮廓点曲率计算

在图像分析和模式识别中,预估数字图像形状的几何特征是重要的步骤,而所有几何特征中,轮廓点曲率特征是决定图像中目标形状的主要特征之一<sup>[13]</sup>.在数字图像中,因为分辨率等问题,轮廓几何形状很少能够被完美的数字化表示,图 2(b)是图 2(a)中红色曲线的局部放大图,他们的不同之处在于分辨率,但该轮廓不光滑,会增加轮廓点曲率计算的难度.为了提高计算的准确率和可行性,曲率计算方法必须把计算轮廓点的局部扰动考虑在内,综合利用小范围内的轮廓点信息,减少并消除噪声的干扰.

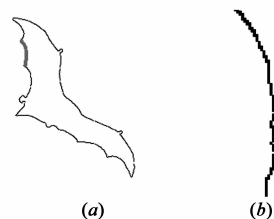


图2 不光滑的数字轮廓

图3是轮廓点  $P_i$  的曲率计算原理图,传统的曲率计算原理为:取相邻的两点  $P_{i-1}$  和  $P_{i+1}$ ,  $P_i, P_{i-1}, P_{i+1}$  三点构成一个三角形并拟合圆周,计算其半径,则轮廓点  $P_i$  的曲率为半径的倒数.但是这种方法没有考虑到  $P_i$  点周围轮廓点的分布情况,且  $P_{i-1}$  和  $P_{i+1}$  可能是噪声点或轮廓不清晰点,这样计算出的曲率值误差较大.为了克服这一影响,综合应用  $P_i$  点相邻附近轮廓点的分布情况,取距离  $P_i$  点  $k$  个距离的两点  $P_{i-k}$  和  $P_{i+k}$ ,  $P_{i-k}, P_i, P_{i+k}$  三点构成一个三角形并拟合圆周,轮廓点的曲率为该圆周半径的倒数.

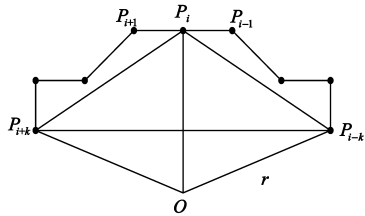


图3 曲率计算原理图

在一个很小的区域内,可以近似地认为  $|P_{i-k}P_i| \approx |P_iP_{i+k}|$ ,根据正弦定理可得:

$$\frac{\sin(\alpha/2)}{r} = \frac{\sin\omega}{|P_iP_{i+k}|} = \frac{\sin(\pi - 2 * \alpha/2)}{(|P_{i-k}P_i| + |P_iP_{i+k}|)/2}$$

$$= \frac{2\sin(\alpha/2)\cos(\alpha/2)}{(|P_{i-k}P_i| + |P_iP_{i+k}|)/2} \quad (0 < \alpha \leq \pi) \quad (1)$$

由公式(1)可得:

$$\frac{1}{r} = \frac{4\sqrt{1 - \sin^2(\alpha/2)}}{|P_{i-k}P_i| + |P_iP_{i+k}|} = K(i) \quad (2)$$

当  $P_{i-k}, P_i, P_{i+k}$  在一条直线上时,  $\alpha = \pi$ ,由公式(2)可得  $K = 0$ ;当  $\alpha$  逐渐减小  $\alpha \rightarrow 0$  时,  $\sin(\alpha/2) = 0$ ,曲率  $K \rightarrow \frac{4}{|P_{i-k}P_i| + |P_iP_{i+k}|}$ .

从图3可以看出:

$$\sin(\alpha/2) = \frac{|P_{i-k}P_{i+k}|/2}{|P_iP_{i+k}|} = \frac{|P_{i-k}P_{i+k}|}{|P_{i-k}P_i| + |P_iP_{i+k}|} \quad (3)$$

将式(3)代入式(2)中可得  $P_i$  点的曲率计算公式为:

$$K(i) = \frac{4\sqrt{(|P_{i-k}P_i| + |P_iP_{i+k}|)^2 - |P_{i-k}P_{i+k}|^2}}{(|P_{i-k}P_i| + |P_iP_{i+k}|)^2} \quad (4)$$

## 2.2 角点检测及轮廓分段

角点是图像轮廓特征中非常重要的特征<sup>[18]</sup>,具有平移、尺度、旋转的不变性.一般认为角点是二维图像灰度变化剧烈的点或图像轮廓曲线上曲率极大值的点.这些点在描述图像轮廓形状特征的同时,有效的减少了处理信息的数据量,提高了计算的速度,有利于目标识别和形状检索.

首先计算轮廓上每一个点处的曲率值  $K(i)$ ,如果  $K(i) \geq T$ ,那么  $P_i$  标定为候选角点.通常,  $k$  值一般选

择为 3-5, 阈值  $T$  的选择与参数  $k$  的选择和轮廓线的弯曲程度有关,由于轮廓点曲率值近似满足正态分布,符合  $3\sigma$  准则,与角点的定义和稀疏性较符合,取  $T = 3\sigma$ .最后通过非极大值抑制方法来获得真正的角点,该算法的计算公式如下:

$$K(i) = \max_{|j-i| \leq k} K(j) \quad (5)$$

即当  $P_i$  的曲率在前后  $k$  个点范围内是最大值时,该点是真正角点,如果该点的曲率小于等于其中其余任何一个点的曲率,那么该点就不是角点.

图4(a)是一副来自 MPEG-7 形状数据库的蝙蝠图像,(b)是应用 Canny 边缘检测算子检测蝙蝠边缘的二值图像,(c)是检测出来的蝙蝠形状所有的候选角点,(d)是采用非极大值抑制算法检测的蝙蝠真正的角点,(e)是按照逆时针方向每个轮廓点的曲率值.从图4(d)可以看出,该角点检测方法能够按照曲率值的原则检测目标轮廓的角点,且效果较好精度很高,误查找率较低,为后续的目标识别工作带来便利.

通过轮廓上角点特征可以将目标形状分成若干段轮廓分段,但是这些轮廓分段并非都是特征完整的,并有利于目标识别,如何将轮廓分段中提取出有效完整的轮廓特征分段是接下来的工作.

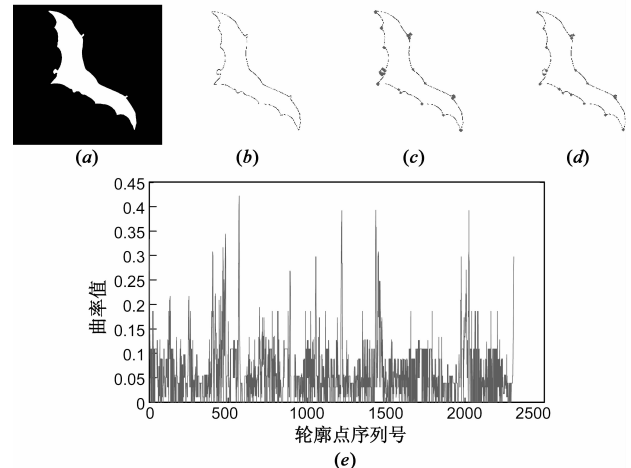


图4 角点检测过程

## 2.3 轮廓特征分段提取

从认知心理学和人类视觉角度出发,凹凸性明显的点和由这些点组成的曲线能够更容易地被人类视觉捕捉到.根据这一准则,本文将轮廓上每一个点按照曲率分布情况分层地描述成特征点和非特征点,轮廓分段按照价值尺度分层地描述成轮廓特征分段和轮廓非特征分段.轮廓特征分段完整高效地描述目标形状特征,既能完整描述目标的形状特征,又能与其他目标区分,识别和检索的准确率更高.

轮廓上特征点和非特征点的描述如下:

$$P_F \{i: K(i) \geq 2\sigma \text{ or } i \in \text{corners}, \sigma = \text{std}(K)\} \quad (6)$$

$$P_{NF} = \{i: K(i) < 2\sigma \text{ and } i \notin \text{corners}, \sigma = \text{std}(K)\} \quad (7)$$

式中,  $K$  代表目标轮廓点曲率值,  $\sigma$  代表所有轮廓点曲率值的标准差,  $\text{std}$  代表标准差。

根据高斯函数  $3\sigma$  法则, 超出  $2\sigma$  范围内的概率不超过 6%, 这符合人类对目标轮廓的认知特点, 即用较少关键的点就可以达到识别目标的目的<sup>[11,14]</sup>。

根据轮廓上点的分布特点, 提出轮廓分段的价值尺度, 定义如下:

$$V(s_i) = N(P_{s_i}^F) / N(P_{s_i}^{NF}) \quad (8)$$

价值尺度是该轮廓分段上特征点数量和非特征点数量的比值, 根据这个定义可以看出, 价值尺度越大, 轮廓分段凹凸性特征越突出, 越容易被识别. 对于整个轮廓, 首先根据角点特征将轮廓分成若干段  $S = \{s_i, i = 1, 2, \dots, n\}$ , 计算每一个轮廓分段的价值尺度, 定义轮廓特征分段的门限阈值:

$$Th_v = \omega_f \max_{s_i \in S} (V(s_i)) \quad (9)$$

式中  $\omega_f$  为轮廓特征分段权重参数,  $\max_{s_i \in S} (V(s_i))$  表示所有轮廓分段中价值尺度最大的数值。

轮廓特征分段和非特征分段的定义如下:

$$S_F = \{i: V(s_i) \geq Th_v\} \quad (10)$$

$$S_{NF} = \{i: V(s_i) < Th_v\} \quad (11)$$

为了将目标轮廓分割成几段特征完整识别率高的特征分段, 提出了一种多级轮廓分段合并算法:

#### 算法 1 多级轮廓分段合并算法

1. 初始化: 给定最小轮廓特征分段数量  $m$  和轮廓特征分段权重参数  $\omega_f$ ;
2. 停止条件: 如果轮廓分段的数量小于  $m$  或者每一个轮廓分段的价值大于门限阈值  $Th_v$ ;
3. 每次合并前选择价值尺度最小的轮廓分段  $s_i$ , 与相邻两个轮廓分段价值尺度最大的那个分段合并, 假设  $s_i^L$  和  $s_i^R$  分别代表  $s_i$  左右两个轮廓分段, 如果  $V(s_i^R) \geq V(s_i^L)$ , 则  $s_i$  与右相邻分段  $s_i^R$  合并, 否则与左相邻分段  $s_i^L$  合并;
4. 合并后的轮廓分段记为  $s'$ , 并从  $S$  中移除分段  $s_i$  和  $s_i^L$  (或  $s_i^R$ ), 将  $s'$  归入到  $S$  中, 计算其价值尺度  $V(s')$ , 如果  $V(s') \geq Th_v$ , 那么选取  $s'$  作为特征分段  $S_F$ .

$\omega_f$  越大,  $m$  越小, 轮廓特征分段的数量越少, 每一个分段特征完整, 有利于识别, 但计算时间较长; 相反,  $\omega_f$  越小,  $m$  越大, 轮廓特征分段数量较多, 计算时间较短, 但并非每一个分段特征都完整, 不利于识别, 需要通过增加更多的细节特征来提高目标识别的可信度. 通过多级轮廓分段合并算法可以将目标图像分割成有限段特征完整的轮廓特征分段, 用于后续的识别。

为了实现该描述算子平移、尺度、旋转的不变性,

将轮廓特征分段进行归一化处理. 将每个特征分段经过等间隔采样后表示为一组有序点集  $\{t_1, t_2, \dots, t_n\}$ , 然后对每个轮廓点按照一定规则进行坐标变换:  $t_1 \rightarrow t_1' = (0, 0)$ ,  $t_n \rightarrow t_n' = (1, 0)$ , 轮廓点  $\{t_2, \dots, t_{n-1}\}$  从原坐标系变换到由  $t_1'$  和  $t_n'$  确定的坐标系中, 完成归一化处理<sup>[12]</sup>. 红色实线的轮廓特征分段归一化过程如图 5 所示:

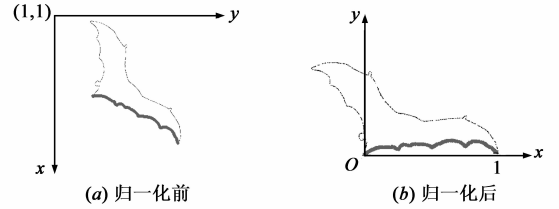


图 5 轮廓特征分段归一化过程

## 3 形状相似度匹配

### 3.1 Shape Context 距离计算

Belongie 等<sup>[2]</sup>提出的 Shape Context (SC) 是最成功的形状描述子, 具有平移、旋转、尺度的不变性, 可以获得目标的全局特征, 且目标的相似度检测的匹配公式能够克服目标的局部变形带来的识别影响, 以下简要回顾该算法的一般原理。

假设目标轮廓  $C$ , 由  $n$  个轮廓点组成  $C = \{x_i, i = 1, 2, \dots, n\}$ . 任意一个轮廓点  $x_i$  可以与其余  $n-1$  个轮廓点构成特征向量形成对数极坐直方图  $h_i$ ,  $h_i$  可以表示为:

$$h_i(k) = \# \{q \neq p_i: (q - p_i) \in \text{bin}(k)\} \quad (12)$$

轮廓点  $p_i$  和  $q_j$  之间的相似度距离定义为:

$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (13)$$

其中  $h_i(k)$  和  $h_j(k)$  分别为  $p_i$  点和  $q_j$  点对应的直方图值,  $K = 60$ .

目标图像  $P$  和  $Q$  间的 Shape Context 距离定义为:

$$D_{sc}(P, Q) = \frac{1}{n} \sum_{p \in P} \arg \min_{q \in Q} C(p, T(q)) + \frac{1}{m} \sum_{q \in Q} \arg \min_{p \in P} C(p, T(q)) \quad (14)$$

其中  $T(\cdot)$  表示经过 TPS 形状转化后的预估值,  $D_{sc}(P, Q)$  越小, 两者之间的相似度越大。

### 3.2 形状相似度检测

当两幅目标图像作匹配计算时, 定义一个距离准则来反映该目标间的相似度. 假设两个形状轮廓  $C_1$  和  $C_2$ , 它们被分成  $n_1$  个和  $n_2$  个轮廓特征分段, 形成特征分段集  $C_{s1}$  和  $C_{s2}$ .

首先, 计算两个目标形状轮廓特征分段间的相似度距离. 对于形状轮廓  $C_1$  和  $C_2$ ,  $C_1$  中任意一个特征分

段  $C_{se_i}$  和  $C_2$  中任意一个特征分段  $C_{se_j}$ , 这两个轮廓特征分段的 SC 相似度距离可以表示为  $d_{ij} = D_{sc}(C_{se_i}, C_{se_j})$ . 对于该两个轮廓中所有特征分段, 两两分段之间的 SC 距离形成的距离矩阵可以表示为:

$$Dis(C_{s_1}, C_{s_2}) = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n_2} \\ d_{21} & d_{22} & \cdots & d_{2n_2} \\ \vdots & \vdots & d_{ij} & \vdots \\ d_{n_1} & d_{n_2} & \cdots & d_{n_1n_2} \end{bmatrix},$$

$$i = 1, 2, \dots, n_1; j = 1, 2, \dots, n_2 \quad (15)$$

其中  $d_{ij}$  表示轮廓  $C_1$  中第  $i$  个特征分段和轮廓  $C_2$  中第  $j$  个特征分段的 Shape Context 相似度距离.

然后, 综合考虑轮廓特征分段长度参数对识别的影响. 从认知心理学角度来看, 人类视觉对目标的大小敏感程度比较大, 轮廓特征分段相对于整个轮廓越大, 则该分段特征越明显, 更易被识别. 两两轮廓分段之间相似度可以表示为:

$$D_{ij} = d_{ij} / \exp(|\alpha_i^l - \alpha_j^l|) \quad (16)$$

其中  $\alpha_i^l = \text{len}(C_{se_i}) / \text{len}(C_1)$ ,  $\alpha_j^l = \text{len}(C_{se_j}) / \text{len}(C_2)$ , 即各自轮廓特征分段的长度与轮廓长度的比值.  $D_{ij}$  越小, 两轮廓特征分段之间的相似度越大.

轮廓特征分段集之间的相似度可以表示为:

$$Sim(C_{s_1}, C_{s_2}) = \begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1n_2} \\ D_{21} & D_{22} & \cdots & D_{2n_2} \\ \vdots & \vdots & D_{ij} & \vdots \\ D_{n_1} & D_{n_2} & \cdots & D_{n_1n_2} \end{bmatrix},$$

$$i = 1, 2, \dots, n_1; j = 1, 2, \dots, n_2 \quad (17)$$

最后, 轮廓  $C_1$  与  $C_2$  之间的相似度可以表示为:

$$Sim(C_1, C_2) = \frac{M \cdot * Sim(C_{s_1}, C_{s_2})}{\text{len}(C_1)\text{len}(C_2)} \quad (18)$$

$$其中 M = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1n_2} \\ M_{21} & M_{22} & \cdots & M_{2n_2} \\ \vdots & \vdots & M_{ij} & \vdots \\ M_{n_1} & M_{n_2} & \cdots & M_{n_1n_2} \end{bmatrix}, \quad \begin{matrix} i = 1, 2, \dots, n_1; \\ j = 1, 2, \dots, n_2, \end{matrix}$$

$$M_{ij} = \begin{cases} 1, & \text{如果 } C_1 \text{ 中特征分段 } i \text{ 与 } C_2 \text{ 中特征分段 } j \text{ 发生匹配} \\ 0, & \text{其它} \end{cases}$$

矩阵  $M$  与  $Sim$  的点积表示矩阵中对应元素的乘积之和.

$Sim(C_1, C_2)$  反映了目标图像  $C_1$  与  $C_2$  两者的相似程度, 该值越小, 则目标图像间越相似.

本文算法的框图如图 6 所示:

## 4 实验和分析

为了论证以上提出方法的精确性, 本文选用 MPEG-7 Shape-1 Part-B 数据库<sup>[15]</sup> 作为测试目标, 并与一些现有的算法作比较. MPEG-7 形状数据库被广泛用来测试形状识别方法, Part-B 主要用来衡量基于相似性方法的检索精度 (Similarity-based retrieval), 是目前衡量形状描述子可靠性的重要指标. 该数据库共有 1400 个形状, 分为 70 类, 每类包含 20 个形状, 在进行检索时, 每个形状都作为待检索目标, 然后在整个数据库中找到最相似的 40 个形状, 如果这 40 个形状中包含了与待检索形状同类的所有 20 个形状, 那么该类的检索正确率为 100%, 统计每个待检索形状的 40 个最佳匹配形状中和待检索形状是同类的形状个数, 将每个形状的正确个数相加除以 2800 (40 \* 70), 即得到了形状描述子在该数据库上的检索精度, 这种度量方法称为 Bull-eye.

图 7 是 2 种目标形状的检索情况, 前 10 个检索的目标形状如下图所示, 奇数行是本文方法检索形状, 偶数行是 SC 算法检索形状. 从图 7 可以看出, 在形状检索率上, 本文提出的算法优于 SC 算子.

由图 7 可以看出, 在形状检索方面, 分层描述目标识别算法优于 SC 形状描述子. 在第二行可以看到一个较惊讶的结果, 基于 SC 描述子, 图像 Hat 检索错了 4 个, 被误认为是 Device3, 因为 SC 描述子只是关注了整个轮廓点的空间位置分布, 而忽略了帽檐局部特征, 而分层描述算法能够提取目标形状的局部特征, 有效抓住了帽檐局部特征变化, 所以前 10 个形状检索都是正确的. 从第四行可以看出, 采用 SC 算法, 图像 Lizard 被检索错了 4 个, 误认为是 Snake, 因为 SC 算法在计算相似度时没有考虑到形状间的长度几何特征, 而本文提出的方法不仅能够提取出 Lizard 脚的形状特征, 还在计算相似度时考虑了轮廓分段的长度特征, 故检索精度本文提出的方法高, 本章方法前十个检索形状都是正确的.

本文方法与其它算法对 MPEG-7 Database 的形状检索率如表 1, 可以看出本文方法是里面检索精度最高的.

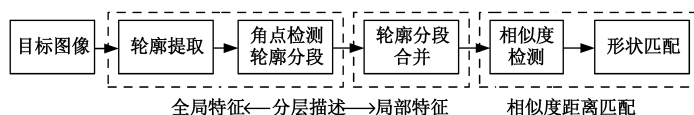


图6 算法框图

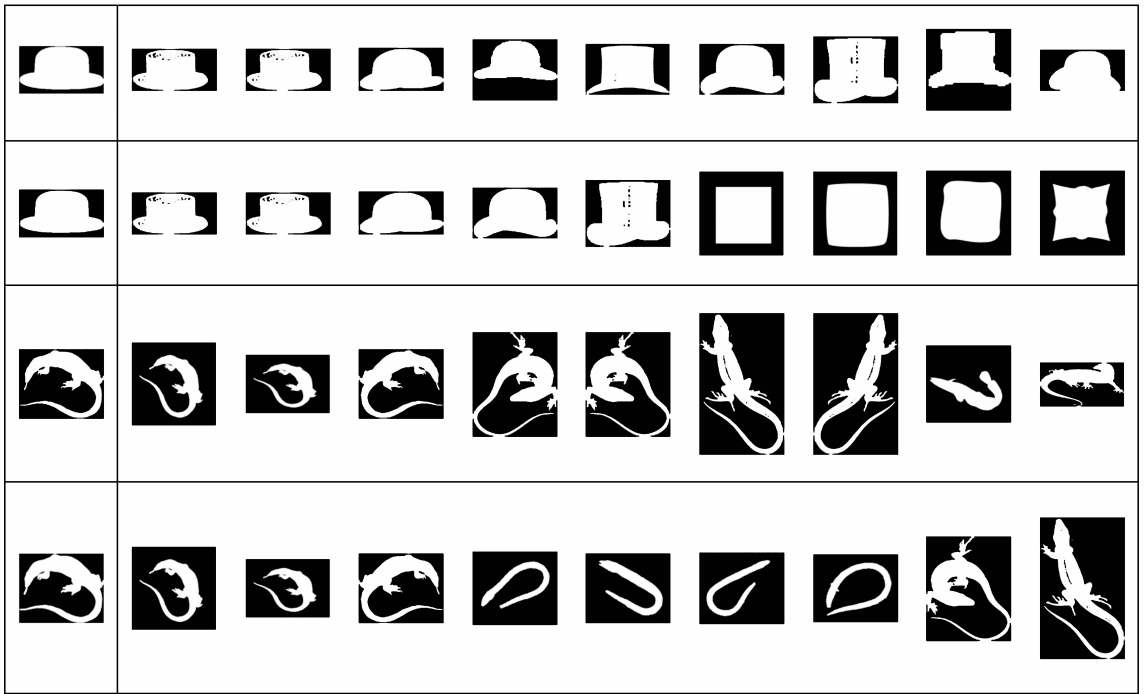


图7 五种目标的检索情况

表 1 MPEG-7 数据库常用算法检索率

Algorithm	Recognition Rate(%)
Shape Tree <sup>[16]</sup>	87.7%
Shape Context <sup>[2]</sup>	76.51%
Inner Distance Shape Context <sup>[8]</sup>	85.4%
Contour Flexibility <sup>[11]</sup>	89.31%
MCC <sup>[17]</sup>	84.93%
Skeletal Context <sup>[18]</sup>	79.92%
Polygonal Shape Description <sup>[19]</sup>	84.33%
Our method	91.05%

为了测试该方法是否对遮挡目标有效,本文从 MPEG-7 数据库选择了 10 幅图像作为测试集,如图 8 所示.

本文设置了当遮挡率分别为 10%, 20% 和 30% 时的目标识别情况,见表 2. 为了保证识别的准确性,避免发生偶然误差,每个遮挡目标都被测试 50 次,平均值作为该目标的识别率.

从表 2 可以看出,本文提出的算法对部分遮挡目标是有效和鲁棒的. 如果目标能够保留主要特征,遮挡率几乎对识别的精度没有影响. Bat, Bird, Camel, Dog, Elephant 和 Deer 的识别率相对较高, Apple, Chicken, Cup, Spoon 的识别率相对较低. 因为前者的局部特征相对多于后者所以虽然部分轮廓丢失了,依然能够被其它的局部特征识别.

将本文方法与其它算法作比较,其识别准确率见表 3. 从表 3 可以看出,本文提出的方法在常用的算法中是最佳的. 本文通过轮廓曲线的曲率分层,提取出了角点这一轮廓重要的特征,根据角点分布将轮廓分段,该分段代表了轮廓的局部特征,对遮挡目标的识别具有较好的鲁棒性. 同时根据价值尺度,将轮廓分段按照多级轮廓分段合并算法合并成有限个代表轮廓主要特征的轮廓特征分段,提高了识别的准确率和鲁棒性. 最后,本文改进了 Shape Contexts 算法,引入了长度尺度,使得轮廓特征分段的相似度计算准确率和可信度更高. 综合以上三个因素考虑,本文提出的分层描述算法能够识别遮挡目标,且效果较好.



图8 MPEG-7中目标测试集

表 2 不同遮挡率下的目标识别率

Occluded Ratio	10%	20%	30%
Apple	92.3	88.2	83.5
Bat	96.5	94.2	93.7
Bird	95.3	93.9	90.4
Camel	94.8	90.6	88.9
Chicken	93.8	90.1	86.4
Cup	93.2	89.7	86.3
Dog	94.8	91.8	89.6
Elephant	95.4	92.3	88.8
Spoon	91.1	88.5	84.3
Deer	93.8	91.6	87.7

表 3 不同遮挡率下的识别准确率

Occluded Ratio	10%	20%	30%
Pol-Des <sup>[20]</sup>	91.4	86.5	80.6
Sub-Mat <sup>[21]</sup>	91.6	87.4	82.1
Seg-Set <sup>[12]</sup>	92.3	89.8	87.9
MCC <sup>[17]</sup>	91.5	86.4	83.8
Con-Sim <sup>[10]</sup>	91.8	86.1	84.9
<b>Our method</b>	<b>92.5</b>	<b>90.6</b>	<b>88.4</b>

## 5 总结和展望

本文提出了一种基于分层描述的形状描述子和综合考虑长度尺度的 Shape Context 相似度距离测度. 该算法通过寻找轮廓角点把轮廓分成若干段轮廓分段, 再根据价值尺度采用多级轮廓分段合并算法得到能够完整高效轮廓形状特征的轮廓特征分段; 将轮廓特征分段之间的 Shape Context 相似度和长度准则结合起来, 有效提高了目标分类识别率和形状检索率, 且对部分遮挡的目标也有较好的识别效果.

在后续研究中, 针对细节特征较多角点检测较难, 或者局部发生变形的目标, 我们将探索如何在保持目标特征基础上简化轮廓, 更高效稳定的描述目标特征和提高算法性能.

### 参考文献

- [1] 许新征, 丁世飞, 史忠植, 等. 图像分割的新理论和新方法[J]. 电子学报, 2010, 38(2A): 76 – 82.  
Xu Xin-zheng, Ding Shi-fei, Shi Zhong-zhi, et al. New theories and methods of image segmentation[J]. Acta Electronica Sinica, 2010, 38(2A): 76 – 82. (in Chinese)
- [2] Serge Belongie, Jitendra Malik, Jan Puzicha. Shape matching

and object recognition using shape contexts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(24): 509 – 522.

- [3] 汤进, 陈展展, 罗斌, 等. 基于复杂网络和最优子序列双射的形状描述与匹配[J]. 电子学报, 2011, 39(8): 1757 – 1765.  
Tang Jin, Chen Zhan-zhan, Luo Bin, et al. Shape descriptor and matching based on complex network and OSB[J]. Acta Electronica Sinica, 2011, 38(8): 1757 – 1765. (in Chinese)
- [4] 王斌. 一种基于多尺度拱高形状描述的图像检索方法[J]. 电子学报, 2013, 41(9): 1821 – 1825.  
Wang Bin. Image retrieval using multi-scale arch height shape description[J]. Acta Electronica Sinica, 2013, 41(9): 1821 – 1825. (in Chinese)
- [5] 黄波, 赵继印, 郑蕊蕊, 等. 基于多尺度自卷积归一化直方图的仿射不变量模式识别[J]. 电子学报, 2011, 39(1): 64 – 69.  
Huang Bo, Zhao Ji-yin, Zheng Rui-rui, et al. Affine invariant pattern recognition based on multi-scale autoconvolution normalized histogram[J]. Acta Electronica Sinica, 2011, 39(1): 64 – 69. (in Chinese)
- [6] 雷阳, 雷英杰, 周创明, 等. 基于直觉模糊核匹配追踪的目标识别方法[J]. 电子学报, 2011, 39(6): 1441 – 1446.  
Lei Yang, Lei Ying-jie, Zhou Chuang-ming, et al. Techniques for target recognition based on intuitionistic fuzzy kernel matching pursuit[J]. Acta Electronica Sinica, 2011, 39(6): 1441 – 1446. (in Chinese)
- [7] 周瑜, 刘俊涛, 白翔. 形状匹配方法研究与展望[J]. 自动化学报, 2012, 38(6): 889 – 909.  
Zhou Yu, Liu Jun-tao, Bai Xiang. Research and perspective on shape matching[J]. Acta Automatica Sinica, 2012, 38(6): 889 – 909. (in Chinese)
- [8] Haibin Ling, David W. Jacobs. Shape classification using the inner-distance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(2): 286 – 299.
- [9] Latecki L J, Lakämper R. Application of planar shape comparison to object retrieval in image databases[J]. Pattern Recognition, 2002, 35(1): 15 – 29.
- [10] Xiang Bai, Xingwei Yang, Longin Jan Latecki. Detection and recognition of contour parts based on shape similarity[J]. Pattern Recognition, 2008, 41(7): 2189 – 2199.
- [11] Chunjing Xu, Jianzhuang Liu, Xiaou Tang. 2D shape matching by contour flexibility[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(1): 180 – 186.
- [12] Kang B Sun, Boaz J Super. Classification of contour shapes using class segment sets[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C]. Washington DC, USA: IEEE, 2005. 727 – 733.
- [13] T. B. Sebastian, P. N. Klein, B. B. Kimia. On aligning curves

- [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(5): 550 – 571.
- [14] Latecki L J, Lakämper R. Convexity rule for shape decomposition based on discrete contour evolution[J]. Computer Vision and Image Understanding, 1999, 73(3): 441 – 454.
- [15] Latecki L J, Lakamper R, Eckhardt T. Shape descriptors for non-rigid shapes with a single closed contour[A]. Computer Vision and Pattern Recognition[C]. Hilton Head Island, SC, USA: IEEE, 2000. 424 – 429.
- [16] Pedro F. Felzenszwalb, Joshua D. Schwatz. Hierarchical matching of deformable shapes[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]. Minneapolis, MN, USA: IEEE, 2007. 1 – 8.
- [17] Tomasz Adamek, Noel E. O'Connor. A multiscale representation method for nonrigid shapes with a single closed contour [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14(5): 742 – 752.
- [18] Euopides G. M. Petrakis, Aristeidis Diplaros, et al. Matching and retrieval of distorted and occluded shapes using dynamic programming[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(11): 1501 – 1516.
- [19] Myung Jin Choi, Antonio, Alan S. Willsky, A tree-based context model for object recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(2): 240 – 252.
- [20] Krolupper F, Flusser J. Polygonal shape description for recognition of partially occluded objects [J]. Pattern Recognition Letter, 2007, 28(9): 1002 – 1011.

- [21] Saber E, Xu Y W, Tekalp A M. Partial shape recognition by submatrix matching for partial matching guided Image Labeling[J]. Pattern Recognition, 2005, 38(10): 1560 – 1573.

### 作者简介



**黄伟国** 男, 1981 年生于安徽休宁, 苏州大学城市轨道交通学院副教授, 研究方向为数字信号处理、数字图像处理、设备状态监测与故障诊断。

E-mail: wghuang@suda.edu.cn



**顾超** 男, 1988 年生于江苏常州, 苏州大学城市轨道交通学院硕士研究生, 研究方向为计算机视觉。



**朱忠奎(通信作者)** 男, 1974 年生于山东梁山, 苏州大学城市轨道交通学院教授, 研究方向为数字信号处理、机械设备状态监测与故障诊断。

E-mail: zhuzhongkui@suda.edu.cn