

一种基于规范化形状的手绘草图检索算法

钱晶,汪增福

(中国科学技术大学自动化系,安徽合肥 230027)

摘要:提出了一种基于仿射变换自适应的形状和骨架特征的手绘草图检索算法.首先采用一种基于傅里叶变换的降维滤波方法对对象进行规范化处理,并以此为基础提取对象轮廓和区域特征;然后提出一种仿射变换自适应骨架提取算法提取对象骨架,得到骨架树特征描述子,并提出一种骨架树匹配算法;最后结合对象轮廓区域和骨架特征来实现草图检索.实验结果表明,该方法对对象仿射变换(平移、缩放、旋转)有很强的适应性,对噪声有较强的鲁棒性.

关键词:轮廓;区域;骨架;仿射变换;草图检索

中图分类号:TP391.41

文献标识码:A

doi:10.3969/j.issn.0253-2778.2010.10.009

A method for freehand sketch retrieval based on affine adaptive skeleton

QIAN Jing, WANG Zengfu

(Department of Automation, University of Science and Technology of China, Hefei 230027, China)

Abstract: A novel freehand sketch retrieval method was presented, which is based on affine transformation adaptive shape and skeleton features. Firstly, a Fourier-based approach to dimensionality reduction and smoothing was employed to process object contour data. And, after normalization, the sampled contour data was used to extract contours and regional features of objects. Then, a new affine adaptive skeletonization algorithm was presented to extract the skeletons of objects and a new skeleton tree descriptor and matching algorithm were derived. At last, a new sketch retrieval method combining contour region and skeleton features was presented. The experimental results show that the method is robust to object affine transformation (translation, scaling, rotation) and noise corruptions.

Key words: contour; region; skeleton; affine transform; sketch retrieval

0 引言

在基于内容的图像检索应用中,仅仅根据提供的示例图像进行检索已远远不能满足用户的需要,开发语义级的图像检索技术已成为迫在眉睫的一项任务.手绘草图作为人类思路外化和交流的一种方式,不仅能表达用户的语义需求,而且自然、直观,已

成为一个新的研究热点.

手绘草图通常用一些有关联的线条来表达一定的语义信息.这些线条合在一起往往构成某种具有不变结构的形状.研究表明,人类视觉能根据目标的形状对物体进行辨别和区分^[1].对基于形状的图像检索技术来说,如何从目标形状抽取反映物体本质特征的确切描述即成为解决问题的关键.目前,用于

收稿日期:2009-05-13;修回日期:2009-09-12

基金项目:国家自然科学基金(60875026)和中国科学院中国新农村信息化研究中心资助.

作者简介:钱晶,男,1980年生,博士生.研究方向:图像检索. E-mail: jackhunt@ustc.edu

通讯作者:汪增福,博士/教授. E-mail: zfwang@ustc.edu.cn

形状描述的主要特征有轮廓特征、区域特征和骨架特征等. 基于轮廓特征的描述方法主要有直线段描述、样条拟合曲线描述、内角直方图描述、高斯参数曲线描述、傅里叶描述子和曲率尺度空间^[2-4]等, 基于区域特征的描述方法主要有区域面积描述、区域纵横比描述、形状无关矩描述和角半径变换 ART^[5]等, 而基于骨架特征的描述方法则包括由骨架最大内切圆半径构造的距离直方图描述^[6-7]等. 以上描述方法存在这样或那样的缺点, 或者因提取的形状信息不完全导致描述过于粗略, 或者对平移、缩放和旋转变换敏感, 或者对噪声敏感, 等等. 鉴于此, 本文提出一种结合轮廓、区域和骨架特征的草图描述方法. 该方法不仅利用了对象外围轮廓信息和内部区域信息, 还充分利用了由骨架表达的形状和拓扑信息, 由此可形成对草图的多特征鲁棒描述. 在此基础上, 通过定义基于轮廓、区域和骨架特征的相似性度量, 构建了一个基于草图的图像检索算法. 上述基于多特征的图像检索算法可在一定程度上克服单纯利用轮廓特征、区域特征和骨架特征的传统方法的缺陷, 提供更为鲁棒的检索结果.

以下, 首先在节 1 中提出一种采用降维滤波方法处理对象轮廓数据并提取对象轮廓和区域特征的方法, 然后在节 2 中提出一种用仿射变换自适应骨架提取算法得到对象骨架树特征描述以及骨架树匹配的方法, 接着在节 3 中提出结合对象形状和骨架特征来实现草图检索的方法并对实验及其结果展开讨论, 最后在节 4 对全文进行总结, 并就未来拟开展的研究进行展望.

1 形状描述符

1.1 对象规范化

为了提取一致性的对象特征, 本文先对对象做规范化处理. 此前, 先提取对象轮廓, 这里的轮廓是指对象边界的外轮廓. 轮廓跟踪方法参见文献[8], 提取出的轮廓为一个点序列. 如图 1 所示, 图 1(b) 为图 1(a) 中对象的轮廓.

1.1.1 轮廓点降维与滤波

获取对象边界外轮廓后, 由于不同对象的轮廓具有不同数量的点, 为了进行匹配, 必须对其以相同数量点进行采样. 显然, 采样点数量的多寡对检索效果及效率有较大影响. 采样点数量越大, 对对象的轮廓描述越精确, 据此对对象间的形状相似性所做的判断越准确, 但相应的计算量也越大; 采样点数量越

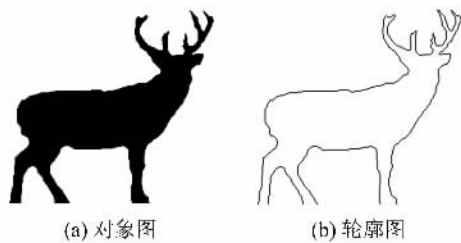


图 1 对象轮廓提取示意图

Fig. 1 Object contour extraction

小, 对象的轮廓细节保持得越少, 也就越平滑, 但对对象的部分形状信息会因此而丢失.

目前主要通过等点采样、等角度采样和等弧长采样等 3 种方法对图像轮廓进行采样以实施降维处理. 设 N 是边界总点数, P 为边界周长, M 为采样点数, 则采样间隔分别为: ①等点采样: N/M ; ②等角度采样: $\theta=2\pi/M$; ③等弧长采样: P/M . 以上 3 种方法对圆形或近似圆形的对象进行采样有较好的效果, 但对复杂轮廓的对象来说效果则不是很理想, 而且还易受噪声干扰. 数字图像中对象的轮廓线常常带有噪声和分割错误, 忽略边界扰动的影响保留原对象视觉感观上的重要成分是对对象识别的要求. 为此, Latecki 等^[9]提出了离散曲线演化模型, 离散曲线的演化实际上是一个轮廓的预处理过程, 这个过程意义主要在两方面: 除去边界噪声, 保留对视觉感官上的重要部分.

基于以上保形去噪的想法, 本文采用一种基于离散傅里叶变换的轮廓点降维方法, 此方法得到的采样轮廓点经过规范化处理后可满足各种不变性要求, 而且有较好的去噪保形效果. 具体做法如下:

(I) 参数化: 首先将对象轮廓参数化得到复信号 $z = \langle z_0, z_1, \dots, z_{N-1} \rangle$, 其中 $z_k = x_k + jy_k$ ($j = \sqrt{-1}$), x_k 和 y_k 分别是对象轮廓上的第 k 点的横、纵坐标.

(II) DFT: 通过离散傅里叶变换(DFT)将信号 z 映射到频域,

$$Z_n = \sum_{k=0}^{N-1} z_k e^{-j\frac{2\pi}{N}kn} = R_n e^{j\theta_n} \quad (1)$$

其中, $n = -N/2, \dots, -1, 0, 1, \dots, N/2$, R_n 和 θ_n 分别表示第 n 个 DFT 系数的振幅和相位.

(III) 降维: 由于对象的大致形状信息集中在低频部分, 所以为了保持对象的形状和消除高频噪声的影响, 我们仅使用 DFT 变换后的低频部分^[10-11], 这里选用频率靠近 0 的 M ($M \ll N$) 个 DFT 系数, 即

令 $m = -M/2, \dots, -1, 0, 1, \dots, M/2 - 1$. M 的取值必须既要能准确表征对象的原始形状,也要保证提取效率.为此,我们考虑数据集的频谱特性,定义 Z 保留 X 个系数后的能量为:

$$E(X) = \sum_{m=-\frac{X}{2}, m \neq 0}^{\frac{X}{2}-1} |Z_m|^2 \quad (2)$$

排除直流成分 (DC component, $m = 0$) 是因为它只提供对象位置信息,对形状描述没有贡献.

图 2 为 MPEG-7 形状库中 Fishes 数据集的 $E(M)/E(N)$ 的均值.由图可知,当 $M \in [16, 128]$,保留的能量与总能量比值为从 83.7% 到 98.4%.

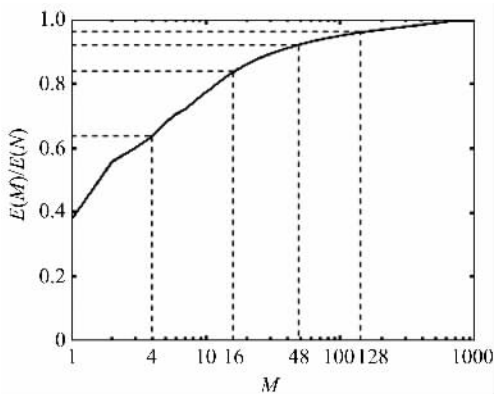


图 2 前 M 个 DFT 系数保留的能量与总能量比值

Fig. 2 The first M DFT coefficients fraction of energy

(IV) 标准化: 为保证平移、缩放、旋转不变性, 和与起点选择无关性, 对 DFT 系数做相应的修改^[12-13], 表 1 显示了标准化 DFT 系数所采取的操作, 标准化后的 DFT 系数 $\hat{Z}_m = \hat{R}_m e^{j\hat{\theta}_m}$ 满足各种不变性要求.

表 1 标准化傅里叶系数

Tab. 1 The Normalized DFT coefficients

不变性	修改后的系数
平移	$\hat{Z}_0 = 0$
缩放	$\hat{R}_m = R_m / R_1$
旋转	$\hat{\theta}_m = \theta_m - (\theta_{-1} + \theta_1) / 2$
起点选择	$\hat{\theta}_m = \theta_m + m(\theta_{-1} - \theta_1) / 2$

(V) IDFT: 对以上标准化 DFT 系数作逆变换 (IDFT), 得满足不变性要求的标准化复信号 \hat{z} ,

$$\hat{z}_k = \frac{1}{M} \sum_{m=-M/2}^{M/2-1} \hat{Z}_m e^{j\frac{2\pi}{M}mk} \quad (3)$$

式中, $k = 0, 1, \dots, M-1$; $\hat{z} = \langle \hat{z}_0, \hat{z}_1, \dots, \hat{z}_{M-1} \rangle$ 即为标准化轮廓采样点序列.

1.1.2 形状规范化

为了使轮廓采样点便于后续处理, 先提取对象区域, 再做如下形状规范化:

(I) 计算形状质心位置 (x_c, y_c) , 平移形状, 从质心位置到坐标原点;

(II) 计算形状主轴方向, 旋转形状 $\Delta\theta$, 从主轴到参照坐标轴;

(III) 缩放形状具有半径 r .

依次连接轮廓采样点即得到转换对象规范化采样轮廓, 提取由轮廓所包围的区域即为转换对象规范化区域.

为了考察转换对象区域是否能有效表示原对象区域, 将原对象和转换对象做同样的形状规范化. 引入区域拟合度来考察两个对象的区域差异, 即比较图 3 中灰色区域和实线所包围区域的差异.

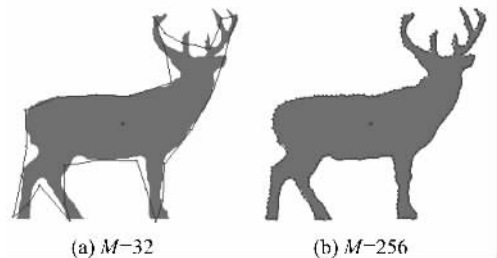


图 3 对象区域差异比较

Fig. 3 The difference of object regions

设 R_0 为原对象区域, R_s 为转换对象区域. A_0 和 A_s 分别为原对象和转换对象区域面积:

$$A_0 = \sum_{(x,y) \in R_0} I_0(x,y), \quad A_s = \sum_{(x,y) \in R_s} I_s(x,y)$$

A_{0s} 和 A_{s0} 定义为:

$$A_{0s} = \sum_{\substack{(x,y) \in R_0 \\ (x,y) \notin R_s}} I_0(x,y), \quad A_{s0} = \sum_{\substack{(x,y) \in R_s \\ (x,y) \notin R_0}} I_s(x,y)$$

其中,

$$I(x,y) = \begin{cases} 1 & \text{if } (x,y) \in R \\ 0 & \text{if } (x,y) \notin R \end{cases}$$

则两个对象区域的拟合度定义为

$$\text{sim } R = 1 - (A_{0s} + A_{s0}) / (A_0 + A_s) \quad (4)$$

这里, 两个对象的区域差异用区域拟合度来反映, 拟合度越大则区域差异越小, 反之则区域差异越大.

对两个对象比较不同轮廓采样率 M 情况下的区域拟合度, 如图 4 所示为 MPEG-7 形状库中 Fishes 数据集的区域拟合度的均值, 可知, 当 $M \in [16, 128]$, 两个对象的区域拟合度为从 81.3% 到

97.6%。因此,当 M 取值大于一定值时,可以用转换对象区域来代替原对象区域。

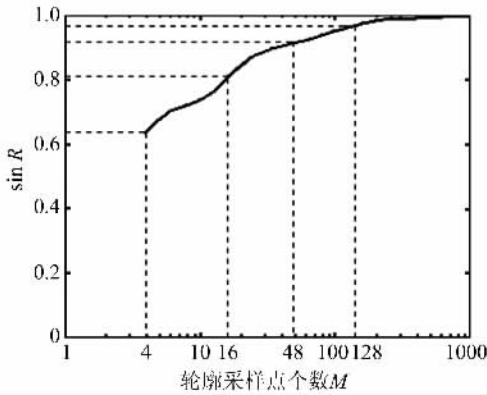


图 4 两个对象的区域拟合度曲线

Fig. 4 The degree of fitting of two object regions

至此,将对象做了规范化处理,规范化后的对象质心在坐标原点(这里为图像中心,坐标为(0,0)),姿态与处理前原对象相同,缩放半径统一为 r ;对象形状基本保持不变,边界平滑,且满足各种不变性要求。

为了便于使用快速傅里叶变换(FFT),通常以

2 的整数次幂的点数采样.综合考虑采样效果和效率,以及后面对象特征提取的需求,在实验的基础上,本系统对每个对象轮廓取 $M=128$ 个采样点。

图 5 显示了一些对象轮廓用上述方法采样并经过规范化处理后的轮廓点示例,此处 $M=128$. 中间的点为对象质心。(b)为(a)的缩放图,(c)为(a)的旋转图,平移前后的采样轮廓是完全一样的。

图 6 演示了对象边界有噪声污染时的采样轮廓.噪声见(b)中 3 个圆圈,(c)中的 3 个圆圈为噪声处的采样轮廓点(黑点所示).对比图 5(a)中的采样轮廓可知,此方法在噪声处用更多的点来进行滤波处理从而平滑了对象边界。

图 7 演示了不同采样点的对象轮廓. M 值分别为 16, 32, 64, 128. 结合图 2 所示 $E(M)/E(N)$,可知,采样点个数 M 在 64 以上时,采样轮廓对原对象轮廓已有很好的近似。

1.2 轮廓描述符的构造

1.2.1 轮廓特征生成

目前描述轮廓的方法有复数坐标、极坐标、中心距离、倾斜角度、累积角度、曲率等.本文采用了质心距离轮廓描述子,即使用对象轮廓点到其质心(x_c ,

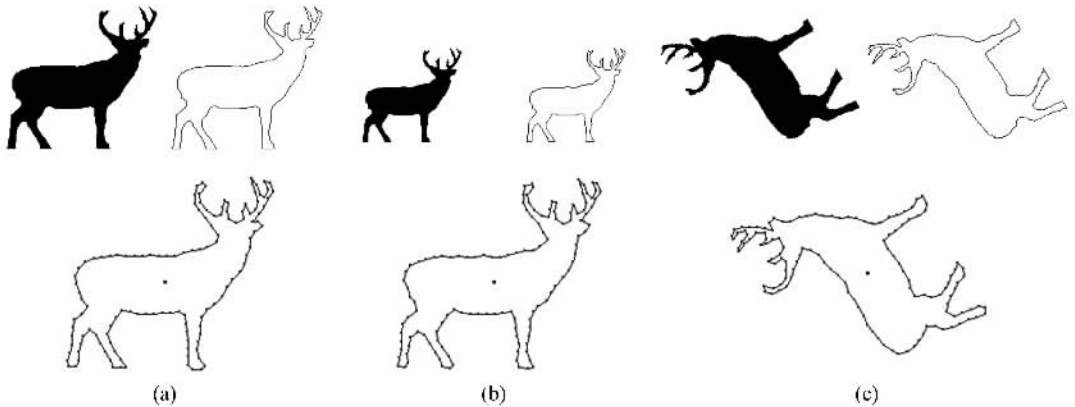


图 5 一些示例对象、对象轮廓和对象规范化轮廓

Fig. 5 Some objects, object contours and their normalized contours

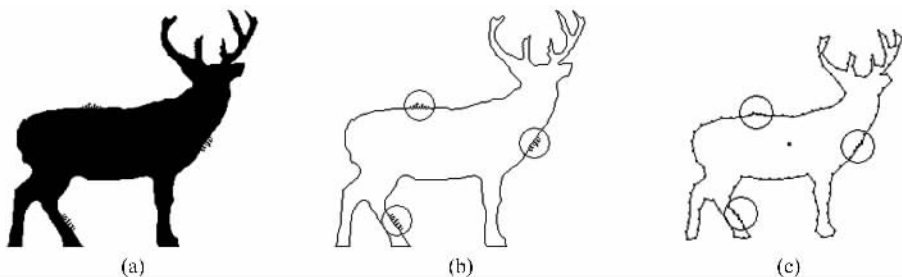


图 6 对象边界有噪声时的规范化轮廓

Fig. 6 Normalized contour of object with noise corruptions

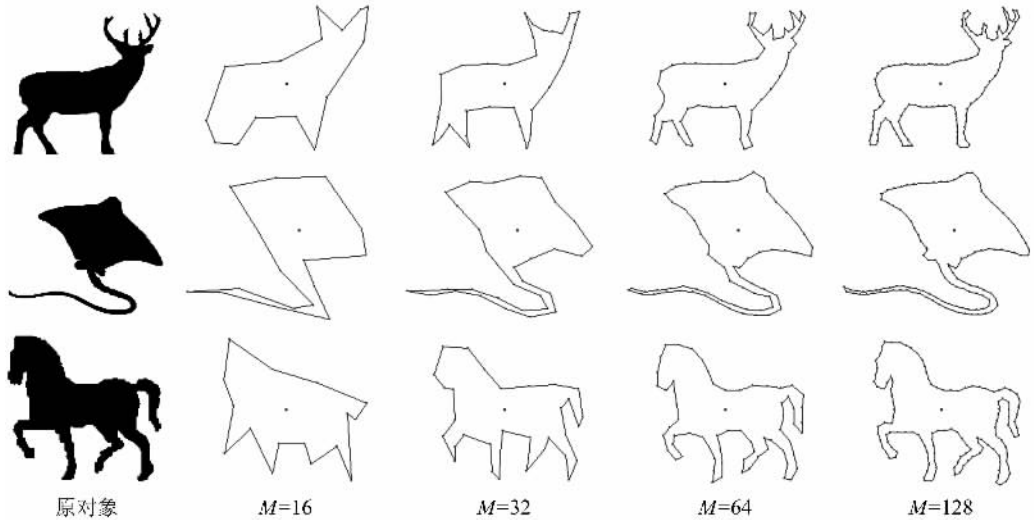


图 7 不同采样点 M 的对象轮廓

Fig. 7 Object contours with different sample rates

y_c) 的距离来表示:

$$r_k = ((x_k - x_c)^2 + (y_k - y_c)^2)^{1/2} \quad (5)$$

式中, $k=0, 1, \dots, M-1$; 由于 (x_k, y_k) 为规范化轮廓采样点, 则 $(x_c, y_c) = (0, 0)$, 且 r_k 满足各种不变性要求. 为了减少噪声影响^[14], 对 r_k 进行离散傅里叶变换:

$$R_m = \frac{1}{M} \sum_{k=0}^{M-1} r_k e^{-j\frac{2\pi k}{M}m} = R'_m e^{j\theta_m} \quad (6)$$

式中, $m=0, 1, \dots, M-1$; R'_m 和 θ_m 分别表示第 m 个 DFT 系数的振幅和相位.

质心距离是实值函数, 因此用 R'_m 代替 R_m 来表征轮廓特征. R_0 称为直流分量, 只提供对象位置信息, 对轮廓描述没有贡献, $R'_0 = R_0$. 由前 $M/2$ 个 R'_m 除以 R'_0 即得到对象的轮廓特征向量:

$$F_C = \left(\frac{R'_1}{R'_0}, \frac{R'_2}{R'_0}, \dots, \frac{R'_{M/2}}{R'_0} \right) = (f_0, f_1, \dots, f_{M/2-1}) \quad (7)$$

1.2.2 轮廓相似性度量

目前, 已经有多种距离表示方法来衡量相似距离, 包括城区距离、欧氏距离、余弦距离、直方图交叉距离、 X^2 统计距离和 Mahalanobis 距离等. 本文采用欧氏距离:

$$D_C(F_Q, F_T) = \left(\frac{2}{M} \sum_{i=0}^{M/2-1} (f_Q - f_T)^2 \right)^{1/2} \quad (8)$$

式中, F_Q 和 F_T 分别为两个对象的轮廓特征向量.

1.3 区域描述符的构造

1.3.1 区域描述子

本文处理对象区域为封闭轮廓所包围区域, 已

做规范化处理. 借鉴 MPEG-7 区域描述符 ART 算法^[5]的思想, 本文将图像平面按径向和角度方向划分成若干子区间, 对每个子区间中的对象区域进行统计, 提取形状特征组成该对象图像的区域描述符向量. 具体做法如下:

(I) 以对象区域 R 的质心点为极坐标原点, 对 R 内的各点进行极坐标变换, 并做归一化处理, 得标准化极坐标 (ρ, θ) , 满足平移、缩放、旋转不变性.

(II) 按照径向和角度方向将极坐标单位圆平面分成 $M \times N$ 个子区间. 统计属于区域 R 的点落在各个子区间的个数, 用 $\{S_R(m, n)\}$ 表示:

$$S_R(m, n) = \sum_{\rho=\frac{m}{M}}^{\frac{m+1}{M}} \sum_{\theta=\frac{n\pi}{N}}^{\frac{(n+1)2\pi}{N}} I(\rho, \theta) \quad (9)$$

式中, $m=0, 1, \dots, M-1, n=0, 1, \dots, N-1$;

$I(\rho, \theta) = \begin{cases} 1 & \text{if } (\rho, \theta) \in R \\ 0 & \text{if } (\rho, \theta) \notin R \end{cases}$. 由于 (ρ, θ) 对各种变换不敏感, 则 $\{S_R(m, n)\}$ 亦对各种变换不敏感.

(III) 对 $\{S_R(m, n)\}$ 中数据分别按径向做一维离散傅里叶变换得到区域描述符 $\{F_R(m, k)\}$.

1.3.2 区域相似性度量

这里采用欧氏距离来进行区域相似性度量:

$$D_R(F_Q, F_T) = \left(\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{k=0}^{N-1} (f_Q(m, k) - f_T(m, k))^2 \right)^{1/2} \quad (10)$$

式中, F_Q 和 F_T 分别为两个对象的区域特征向量. 本文取 $M=4, N=16$.

2 骨架描述符

“骨架”(Skeleton)是图像形态学中的一个概念,是二维对象的一维中线表示.图像中对象的骨架(或中轴)是指包含在对象内部的最大内切圆的圆心轨迹或“烧草”变换过程中的灭火线.现有的骨架提取算法主要有三类:一类是细化算法^[15-16],第二类是中轴变换法^[17-18],第三类是 Voronoi 算法^[19].参照 Zhu 等^[20]的骨架建模方法,本文提出了一种仿射变换自适应的骨架提取算法,并提出一种基于骨架树的等级骨架匹配方法对骨架模型进行相似度匹配.

2.1 骨架基元定义

动植物的骨架存在一定的共性,并且可以解析为许多结构类似的小单元.由此可以设想这些骨架是由几种基本单元(简称为基元)以及它们的变形组成的.基于这种思想,本文首先定义两种骨架基元:骨架点基元和骨架分支基元.

骨架点基元定义:骨架线上的点,这里为对象轮廓内部的最大圆心点,骨架点是构成骨架的最小元素.骨架点类型分为三种:端点,仅有 1 个邻接点的骨架点,即骨架分支末端的点,其八邻域骨架点个数和为 1,端点对应着区域轮廓的突出部分;分叉点,有 3 个或 3 个以上邻接点的骨架点,由此点分出多个分支,其八邻域骨架点个数和大于等于 3;中间点,有 2 个邻接点的骨架点,骨架上除端点和分叉点外的点都为中间点,其八邻域骨架点个数和等于 2.把端点和分叉点统称关键点.

骨架分支基元:若 2 个骨架点之间仅含有中间点则称这 2 个点为直接相连的.直接相连的 2 个关键点之间的所有点的集合称为骨架分支.骨架分支分为两种:内分支,分支的两个关键点都是分叉点;外分支,分支的两个关键点至少有一个是端点.

2.2 仿射变换自适应骨架提取算法

性能良好的骨架化算法应能满足以下条件:保持原始形状的拓扑属性不变,对目标边界噪声不敏感,能够提取出发生在多尺度上的骨架,骨架具有单像素宽度.针对以上要求,本文提出一种仿射变换自适应的骨架提取算法,主要有以下三个步骤:

(I) 对象规范化及预处理

先用前文所述基于离散傅里叶变换的轮廓点降维方法提取对象轮廓并做规范化;再提取规范化对象区域.如前所述,得到的对象区域保持原对象形

状,有统一的尺度,对噪声有一定过滤,且满足各种不变性要求.做预处理使输入图像变为二值图像,预处理之后图像 $I'(x, y)$ 为黑白两区域, $I'(x, y) = 1 - I(x, y)$, $I'(x, y) = 0$ 表示点在黑色区域为对象, $I'(x, y) = 1$ 表示点在白色区域为背景.

(II) 骨架点计算

图像处理中的骨架实际上对应为对象轮廓内部的中轴线.因此本文采用的方法思想是设想用一个可变形且始终保持与对象轮廓相切的圆滚过对象内部,如图 8 所示,那么圆心的轨迹就是骨架线.

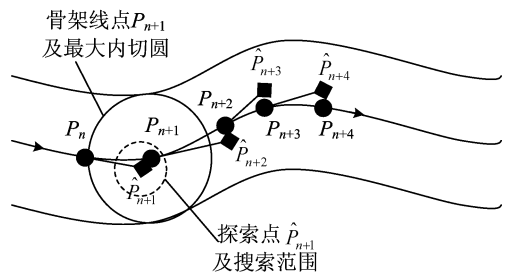


图 8 骨架上点 $P(x, y)$ 的确定

Fig. 8 Computation of skeleton point $P(x, y)$

求骨架点 $P(x, y)$ 的具体算法如下:

①初始化,设定旋转角度差 $\Delta\theta = 360^\circ/m$ 、移动步长 sL 、搜索范围 sR ,确定初始搜索点 (x_0, y_0) 及起始搜索方向 $st\theta$;初始参数根据规范化对象尺度确定.

②对以搜索点 (x, y) 为中心 sR 为半径的圆形区域内的所有点(如图 8 中的虚线小圆所示),寻找对象轮廓内半径最大的内切圆,并返回圆心坐标 (x', y') 和半径 r ,令 $x = x', y = y'$,此圆心点为骨架线点,设为 $P_n(x, y)$.

③以当前骨架线点 $P_n(x, y)$ 为中心,以 $st\theta$ 为起始方向,按逆时针方向旋转,每隔 $\Delta\theta$ 向区域轮廓作射线,设每条射线的长度为 d_i ,它与起始方向 $st\theta$ 的夹角为 $\theta_i = i\Delta\theta$.统计这 m 条射线的长度,找出局部极值的射线.若只有一个局部极值 d_i ,则 θ_i 即为当前搜索方向;若有多个局部极值,则分别以局部极值的方向为搜索方向,并记点 $P_n(x, y)$ 为分叉点.

④令当前点为 $P_n(x_n, y_n)$ 、当前搜索方向为 θ_n ,沿搜索方向移动步长 sL 使 $P_n(x_n, y_n)$ 至 $\hat{P}_{n+1}(x_{n+1}, y_{n+1}) = (x_n + sL \cos \theta_n, y_n + sL \sin \theta_n)$ 处,若 $I'(x_{n+1}, y_{n+1})$ 为 1,则表示 \hat{P}_{n+1} 已在对象轮廓外或超出图像边界,结束该骨架分支搜索.否则,以 $\hat{P}_{n+1}(x_{n+1}, y_{n+1})$ 点为搜索点执行②找到点 P_{n+1} .若

$P_n P_{n+1}$ 连线与已存在骨架线相交或有点在区域外则 P_{n+1} 为非有效点, 若 P_{n+1} 为有效点, 则连接 $P_n P_{n+1}$ 为新增骨架线, 并令 P_{n+1} 为当前骨架线点.

⑤ 返回③继续搜索, 直至所有骨架分支搜索完毕, 即得到所求对象骨架.

(III) 删除冗余分支线

由于对象轮廓的复杂性, 由第(II)步所得骨架线含有若干冗余分支, 故本文通过以下方法删除冗余分支线:

① 对端点, 若其前一点与之延长线在区域内与骨架线相交, 则删除此分支线到其分叉点, 并更新分叉点属性.

② 若端点到其分叉点的线长小于某阈值, 则删除此分支线, 并更新分叉点属性.

③ 重复①, ②直到删除所有不符合要求的分支线.

图 9 为本文算法的骨架提取结果. (a) 为仿射变换和边界噪声情况下的对象“鹿”骨架, 由图可知, 经过形状规范化后提取出的骨架基本一样; (b) 演示了不同采样点 M 值的对象骨架. 由实验结果可知, 本文骨架提取算法有较好的效果.

2.3 骨架描述子生成

2.3.1 骨架树定义

为了突出对象的拓扑特征, 我们把骨架映射到

一个树状结构中, 称之为骨架树.

骨架树的根节点为骨架主轴, 定义为骨架中具有较大辐射域和方向一致性的分支基元依次连接而成的分支链(branch chain, BC). 骨架主轴的构成包括两个要素: 辐射域和方向, 它在整体上刻画了骨架的主要形态和全局结构信息. 后面将给出分支辐射域和方向一致性的具体定义.

骨架树中的孩子节点为骨架分支, 其中叶子节点为骨架外分支, 分支节点为骨架内分支.

骨架树中的边表示分支相邻关系.

2.3.2 骨架基元描述

在提取骨架特征之前, 先定义两种骨架基本描述: 骨架分支描述和骨架关键点描述, 分别用长条模型和圆盘模型来表示. 设定提取的对象骨架均由这两种基元及其变形组成, 它们均可以由一组形状变量表示.

长条模型由一个和两个圆相连的可变形长方形组成(如图 10(a)所示), 圆心 C_1 和 C_2 为骨架结构中的关键点. 我们将中轴线 $C_1 C_2$ 等分成 n 段, 记等分点位置为向量 $\mathbf{X} = (x_1, y_1, \dots, x_n, y_n)$, 在每个等分点处记下中轴线的垂线段长度 $\mathbf{R} = (r_1, r_2, \dots, r_n)$, 用来反映长条模型的形状. 由此定义长条模型的形状向量 $\mathbf{W} = (l, \mathbf{X}, \mathbf{R})$, l 为中轴线长度. 长条模型的变形如图 10(c)所示, 这里中轴线的垂线段长

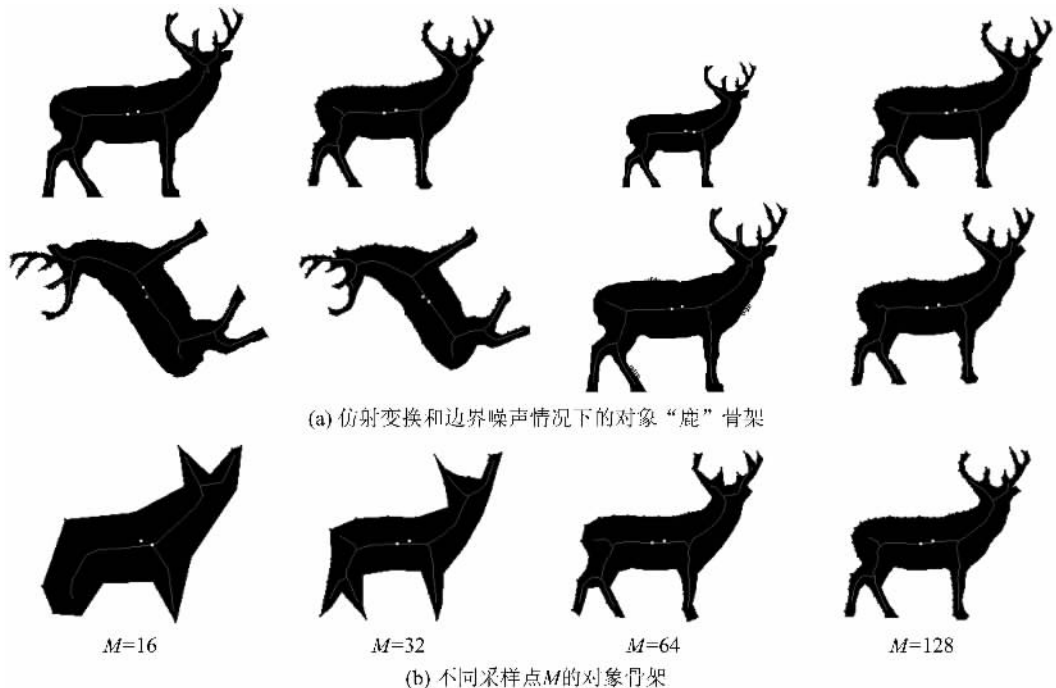


图 9 对象骨架提取结果

Fig. 9 Some results of object skeleton extraction

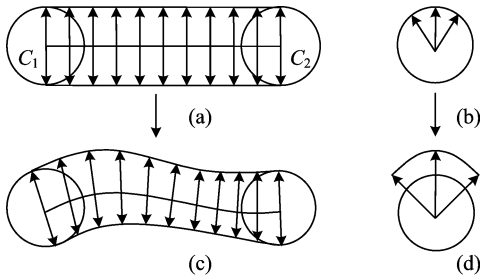


图 10 两种骨架基元: (a) 为长条模型, (b) 为圆盘模型, (c) 和 (d) 为发生变形后的长条模型和圆盘模型
Fig. 10 Two skeleton primitives and their deformations

度 R 用每个等分点为圆心的最大内切圆半径来表示, 做归一化处理, 则变形基元长条模型形状向量可记为

$$\begin{aligned} W' &= (\delta_l l, X'(\delta_\rho \rho_1, \delta_\theta \theta_1, \delta_\rho \rho_2, \delta_\theta \theta_2, \dots), \\ &R'(\delta_r r_1, \delta_r r_2, \dots)) = \\ &(l', X'(\rho'_1, \theta'_1, \rho'_2, \theta'_2, \dots), R'(r'_1, r'_2, \dots)) \end{aligned} \quad (11)$$

式中, $\delta_l = 1/l_{\max}$, $\delta_\rho = 1/\rho_{\max}$, $\delta_\theta = 1/2\pi$, $\delta_r = 1/r_{\max}$, 均为归一化系数; (ρ_i, θ_i) 为等分点相对于对象质心点的极坐标。

圆盘模型为一个可形变的圆(图 10(b)), 我们从圆心出发, 每隔 $360^\circ/(m-1)$ 夹角做一条到轮廓的射线. 记这组射线长度向量为 $D = (d_1, d_2, \dots, d_{m-1})$, d_i 为第 i 条射线段的长度 ($i=1, \dots, m-1$). 由此得圆盘模型形状变量 $C = (x, y, D)$, (x, y) 为圆心位置. 圆盘模型的变形如图 10(d) 所示, 做归一化处理, 则变形基元圆盘模型形状向量可记为

$$\begin{aligned} C' &= (\delta_\rho \rho, \delta_\theta \theta, D'(\delta_d d_1, \delta_d d_2, \dots)) = \\ &(\rho', \theta', D'(d'_1, d'_2, \dots)) \end{aligned} \quad (12)$$

式中, $\delta_\rho = 1/\rho_{\max}$, $\delta_\theta = 1/2\pi$, $\delta_d = 1/d_{\max}$, 均为归一化系数; (ρ, θ) 为圆盘圆心点相对于对象质心点的极坐标。

2.3.3 骨架主轴提取

设 T_B 为骨架分支基元的集合, 对每个分支 $b_k = uv \in T_B$, u, v 为骨架分支的关键点基元, 基于以上骨架分支描述, 分支辐射域 A_R 定义为

$$A_R = \sum_{i=1}^n 2r_i l_i, L = \sum_{i=1}^n l_i \quad (13)$$

式中, n 为骨架分支分成的段数, l_i 为第 i 段骨架线长。

找到内切圆半径最大且靠近对象质心的骨架点. 若此点为分叉点, 则包含此点的分支辐射域最大

的分支为关键分支; 若此点为非分叉点, 则所在分支设为关键分支。

设分支 $b_k = uv, b_j = vw$ 相邻, b_j 相对于 b_k 的方向定义为 uv 延长线和 vw 的夹角 $\theta_{jk} \in [0, \pi]$ 。

则骨架主轴可由以下步骤提取:

(I) 找到所有包含关键分支的分支链, 分支链的两个端分支均为骨架外分支。

(II) 计算分支链辐射域:

$$A_{RC} = \sum_{k \in \Gamma_C} \sum_{i=1}^{n_k} 2r_{ki} l_{ki} \quad (14)$$

式中, Γ_C 为分支链中所含分支序号集合。

(III) 计算分支链方向一致性, 用其方向累积差表示:

$$\theta_C = \left(\sum_{\substack{j, k \in \Gamma_C \\ j, k \text{ 相邻}}} \theta_{jk} \right) / (N_C - 1) \quad (15)$$

式中, N_C 为分支链中分支个数。

(IV) 定义分支链明显度为

$$\text{Vis}_C = \omega A_{RC} + (1 - \omega)(1 - \theta_C / \pi) \delta_\theta \quad (16)$$

式中, δ_θ 为调节系数. 则骨架主轴为明显度最大的分支链。

如图 11 所示骨架中, 分支(1)为关键分支, 分支(1 2 4 7 9 10)所形成的分支链为骨架主轴。

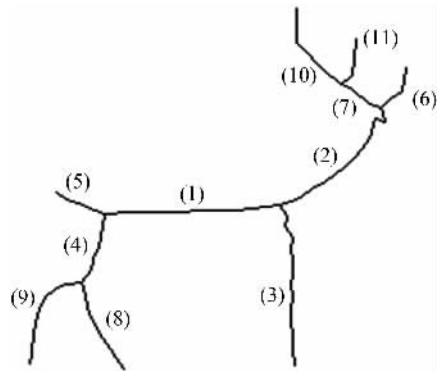


图 11 对象“鹿”的骨架图

Fig. 11 The skeleton of the object “deer”

2.3.4 骨架特征树的生成

根据骨架相邻信息提取骨架分支邻接图. 令骨架主轴为树的根节点, 由分支邻接图得到如图 12 所示骨架树, 节点内“_”前数字表示分支所在树的层次, “_”后数字为节点序号, “(·)”中数字为骨架分支序号, 树的边表示分支相邻关系。

因此, 对象骨架特征可以表示为一个树, 节点为各分支用以上两种模型的归一化描述, 每个节点赋予相应权重, 树边表示成骨架分支邻接图。

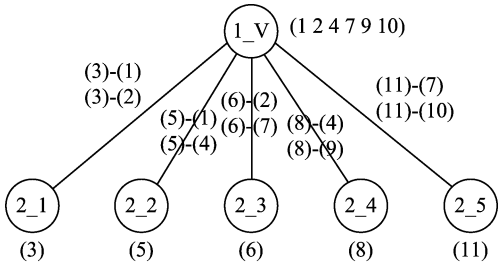


图 12 骨架树

Fig. 12 The skeleton tree

根据骨架树中节点的两模型描述 $W=(L, X, R)$ 和 $C=(x, y, D)$, 可以大体复现出原对象. 图 13 是通过骨架特征复现出来的对象, 由图可知, 骨架特征能很好地表现原对象的形状和拓扑信息.

2.4 骨架相似性度量

2.4.1 骨架分支相似性度量

设 $b1=W_1(l_1, X_1(\rho_{11}, \theta_{11}, \dots, \rho_{1n}, \theta_{1n}), R_1(r_{11}, r_{12}, \dots, r_{1n}))$ 为骨架树 $T1$ 中的一个长条基元, 它对应于 $T2$ 中的基元 $b2=W_2(l_2, X_2(\rho_{21}, \theta_{21}, \dots, \rho_{2n}, \theta_{2n}), R_2(r_{21}, r_{22}, \dots, r_{2n}))$, 定义分支 $b1, b2$ 相似性距离为

$$D_B(b1, b2) = \exp\left\{1 - 1/\left[\frac{\omega_p}{n} \sum_{i=1}^n (|\rho_{1i} - \rho_{2i}| + |\theta_{1i} - \theta_{2i}|) + \frac{\omega_r}{n} \sum_{i=1}^n |r_{1i} - r_{2i}| + \omega_l |l_1 - l_2|\right]\right\} \quad (17)$$

式中, w 为权重, $\omega_p + \omega_r + \omega_l = 1$. 在本文试验中, 根据分支辐射域的定义和对对象复现的贡献程度, 经验地选取 $\omega_p = 0.4, \omega_r = 0.5, \omega_l = 0.1$.

2.4.2 对象骨架树相似度计算

为了计算 2 个骨架树的相似度, 需要先建立树节点即骨架分支之间的匹配关系, 计算匹配的骨架分支的相似度之和, 以此作为整个骨架的相似度. 设需计算相似度的骨架树为 $T1, T2$, 分别有 k, m 个节点, 不妨设 $k \leq m$, 则 2 个骨架树的相似距离计算如下:

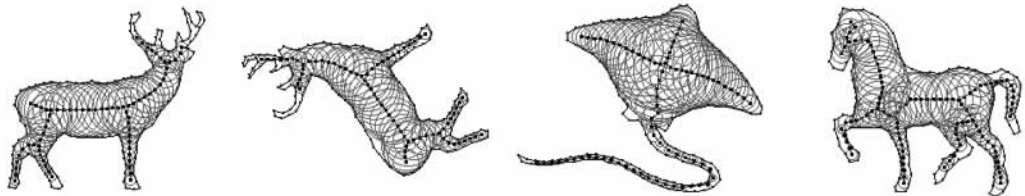


图 13 通过骨架特征复现的对象

Fig. 13 Reproducing objects by skeleton feature

(I) 首先把根节点设为起始匹配节点, 根据其所含分支计算相似性.

(II) 对于孩子节点, 通过迭代的搜索 $T1$ 和 $T2$ 中的匹配节点来得到树的最优匹配. 孩子节点匹配考虑两个因素: 分支拓扑结构和分支相似距离. 分支相似距离由式(17)计算, 定义分支拓扑结构差异为

$$\tau(b1, b2) = \|N(b1) - N(b2)\|,$$

$N(\cdot)$ 为节点的邻接图行向量, 归一化得 $\tau' = \tau/\tau_{max}$. 则匹配节点的相似距离为

$$D_{B_{mi}}(b1, b2) = D_B(b1, b2)\tau(b1, b2) \quad (18)$$

(III) 骨架树的相似距离为

$$D_S(T1, T2) = \frac{1}{m} \left(\sum_{i=1}^k \omega_{B_i} D_{B_{mi}}(b1_i, b2_i) + \sum_{j=k+1}^m \omega_{B_j} D_{B_{max}} \right) \quad (19)$$

式中, $D_{B_{max}} = \max(D_{B_{mi}}(b1_i, b2_i))$, 是对骨架中部分缺失的相似距离表示; ω_{B_i} 为骨架树中节点的重要程度. 显然根节点即骨架主轴的重要程度最大, 因而把 ω_{B_1} 设为 1, 表示骨架主轴对于骨架相似距离的贡献程度最大, 其他分支的重要程度根据所在树的层次和分支特征来计算.

3 手绘草图检索

3.1 草图预处理

用户检索时提交的手绘草图, 由简单的笔划组成, 需要经过预处理, 以得到连续的草图边界外轮廓. 对于用户提交的草图, 采用一种基于八连通域的方向自适应跟踪算法来对草图进行轮廓提取.

图 14 显示了一个手绘草图的处理过程和骨架特征复现出来的对象, (a) 为手绘草图, (b) 为草图轮廓, (c) 为草图规范化采样轮廓, (d) 为草图对象骨架, (e) 为用骨架特征复现出来的草图对象.

3.2 对象相似性度量

物体间形状的“相似”有不同的概念. 如图 15

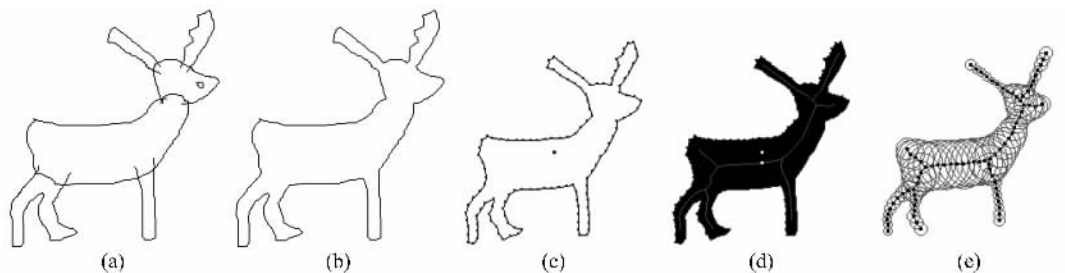


图 14 手绘草图特征获取过程演示

Fig. 14 Feature extraction process of freehand sketch

所示,第一行的各个物体有着相似的空间分布,运用基于区域的相似性准则来观察,它们之间是相

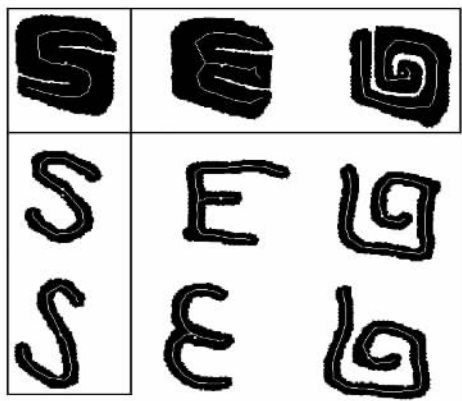


图 15 区域相似、轮廓相似和骨架相似

Fig. 15 Region, contour and skeleton similarity

似的;从轮廓和骨架相似性准则的观点来看,则在每一列上的物体有着较大的相似性;而根据本文骨架特征,每一列的第二和第三个物体具有更大的相似性。

根据前文得到的对象轮廓、区域、骨架的描述子及相似度量,见式(8),(10),(19),则本文用以下距离公式来衡量两个对象的相似性:

$$D = w_C D_C + w_R D_R + w_S D_S \quad (20)$$

式中, D_C, D_R, D_S 分别为对象轮廓、区域、骨架的相似距离; $w_C + w_R + w_S = 1$ 为相应的权重. 根据经验选取 w_C 为 0.45, w_R 为 0.25, w_S 为 0.3.

3.3 实验结果与分析

3.3.1 图像库

本文所采用的图像对象,一部分是从 corel 库图像中提取的对象,一部分是 MPEG-7 形状库中的对象,一部分是用户手绘的草图对象. 图像库中一部分图像对象做了以下仿射变换:平移一定距离图,缩放一定大小图(50%到 150%),逆时针旋转一定角度图(0° 到 360° 的随机角度),受一定噪声污染图. 经

过上述变换后,我们得到的系统图库总共包含 204 幅图像.

3.3.2 性能分析

为了定量评价图像检索算法,需要一种有效的性能评价方法. 目前多种性能评价方法通常是基于主观试验的统计数据,本文使用查准率(precision)和查全率(recall)进行算法效果的评价,定义如下:

$$P = C_N / N, R = C_N / M \quad (21)$$

式中, C_N 表示检索到的相关对象数, N 表示所有检索到的对象数, M 表示图库中相关对象的总数.

查准率反映检索结果的准确度,而查全率反映从图库中检索到相关图像的能力. 查全率和查准率是负相关的,即通常情况下,当查全率提高时,查准率则相应下降. 图 16 是手绘语义对象“鹿”的检索示例,可知不但形状相似的对象,而且经过仿射变换和噪声污染的对象均可被检索出来. 图 17 是对上述图库进行测试的查准率-查全率($P-R$)曲线,其中圆形标记曲线采用轮廓特征,三角标记曲线采用区域特征,菱形标记曲线采用骨架特征,方形标记曲线采用公式(20)三种特征;结合图 15 可知,采用三种特征结合的方法可以更准确地查找出目标图像.

4 结论

本文提出了一种具有仿射变换自适应的形状骨架描述方法和手绘草图检索算法. 该方法引入了一种傅里叶降维滤波方法处理对象轮廓数据并做规范化,以此为基础,先构造轮廓、区域描述符,再提出一种仿射变换自适应骨架提取算法来提取对象骨架、构造骨架树描述符,并提出一种骨架树匹配算法,最后提出结合对象轮廓、区域和骨架特征来实现手绘草图检索. 该方法提取的特征能较好地表现对象,对对象平移、缩放、旋转等变换有很强的适应性,对噪声有较强的鲁棒性. 如何更好地描述对象形状,如何

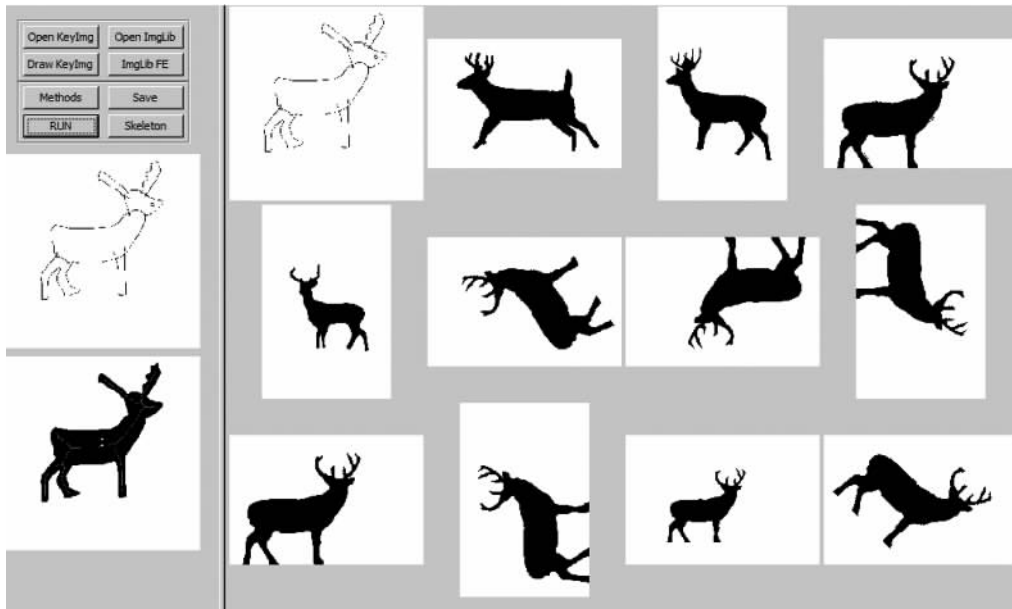


图 16 手绘“鹿”的检索示例

Fig. 16 Retrieval example of sketch object “deer”

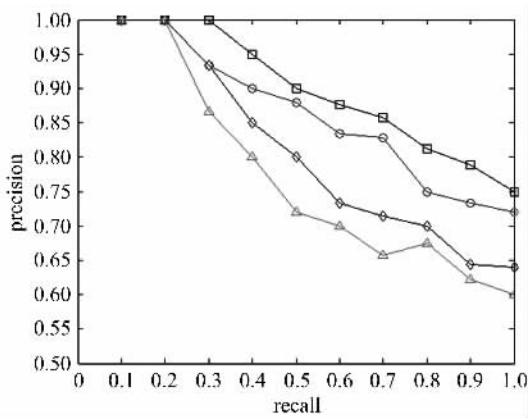


图 17 手绘“鹿”的查准率-查全率曲线图

Fig. 17 Precision and recall of sketch object retrieval

更准确地提取对象骨架以及得到更优的骨架匹配策略将是我们以后研究工作的方向。

参考文献(References)

[1] Zhang D S, Lu G J. Study and evaluation of different Fourier methods for image retrieval [J]. Image and Vision Computing, 2005, 23 (1): 33-49.
 [2] Mokhtarian F, Bober M. Curvature Scale Space Representation: Theory, Applications and MPEG-7 Standardization [M]. Boston, USA: Kluwer Academic Publishers, 2003.
 [3] Makharian F, Abbasi S. Shape similarity retrieval under affine transforms [J]. Pattern Recognition, 2002, 35 (1): 31-41.

[4] Pereira F, Koenen R. MPEG-7: A standard for multimedia content description [J]. International Journal of Image and Graphics, 2001, 1(3): 527-546.
 [5] Bober M. MPEG-7 visual shape descriptors [J]. IEEE Transaction on Circuits and Systems for Video Technology, 2001, 11(6): 716-719.
 [6] Choi W P, Lam K M, Siu W C. Maximal disk based histogram for shape retrieval [C] // Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Hong Kong, China: IEEE Press, 2003, 3: 9-12.
 [7] Choi W P, Lam K M, Siu W C. Extraction of the Euclidean skeleton based on a connectivity criterion [J]. Pattern Recognition, 2003, 36 (3): 721-729.
 [8] Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision [M]. 3rd ed. Belmont, USA: Cengage Engineering, 2007.
 [9] Latecki L J, Lakämper R. Shape similarity measure based on correspondence of visual parts [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2000, 22(10): 1 185-1 190.
 [10] Rafiei D, Mendelzon A O. Efficient retrieval of similar shapes [J]. The International Journal on Very Large Data Bases, 2002, 11(1): 17-27.
 [11] Zhang D S, Lu G J. A comparative study of fourier descriptors for shape representation and retrieval [C] // Proceedings of the 5th Asian Conference on Computer Vision. Melbourne, Australia: IEEE Press, 2002: 646-651.