

基于骨架接合节点特征的形状识别与分类

陈 珺^{1,2} 王 波¹ 周 瑜¹ 白 翔¹ 刘文予¹

(华中科技大学电信系 武汉 430074)¹ (中国地质大学(武汉) 武汉 430074)²

摘 要 介绍了一种利用形状的骨架信息进行形状识别和分类的方法。利用从形状中提取的骨架接合节点信息,将形状相似性问题转化为骨架接合节点距离的计算。采用神经网络的方法,根据输入形状的骨架接合节点距离判断出形状所属类别。该方法可以很好地处理非刚性物体形状识别问题,与现有方法相比,具有更好的形状识别效果。

关键词 骨架,神经网络,形状识别,形状分类

Shape Recognition and Classification Based on Skeleton Junction Nodes Features

CHEN Jun^{1,2} WANG Bo¹ ZHOU Yu¹ BAI Xiang¹ LIU Wen-yu¹

(Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)¹

(China University of Geosciences, Wuhan 430074, China)²

Abstract This paper proposed an algorithm using skeleton information to do shape recognition and classification. Using the information of skeleton junction nodes, the problem of shape similarity is converted to distance of skeleton junction nodes. Using method of neural network, the classification results can be got at the output of the network. The method can solve the recognition of nonrigid objects. The experiment results show that our method has better recognition results than other method.

Keywords Skeleton, Neural network, Shape recognition, Shape classification

1 引言

形状的识别和分类一直是计算机视觉中的一个热点问题,具有广阔的应用前景。而基于骨架的形状表示方式集成了物体的几何和拓扑特性,可以克服基于轮廓的形状表示方式在处理非刚性物体时的障碍,因此可以用在非刚性物体的分类检索中。本文采用神经网络的方法,对将要进行形状分类的骨架特征加以分析,判断出所属的形状。通过在 Tari 数据库^[1]上测试,得到了很好的分类结果。文中的骨架特征是采用 DCE^[2]方法修剪得到的,通过该方法得到的骨架移除多余的不必要分支,完好地保存了物体的结构信息。

2 骨架的相似性算法

图1是同类物体的骨架图。这里骨架的接合节点被标记为a,b,c,d,e,骨架的端节点被标记为1,2,3,4,5,6,7,这些端节点和接合节点就是描述骨架特征的关键点。从图1中可以看到,对于同类物体,其骨架图关键点也可能不同,这为骨架的关键点匹配带来了困难^[3]。本文提出了一种新的骨架图匹配算法,该方法利用最佳序列双射(OSB)^[4]方法解决了两个长度不同的序列匹配问题。这种方法可以跳过两序列中的异常元素,使序列得到较好的匹配关系。经过非相似度计算

得到骨架路径间的距离后,对接合节点序列进行匹配、非相似度计算,再采用神经网络方法训练出一个所需的分类器。

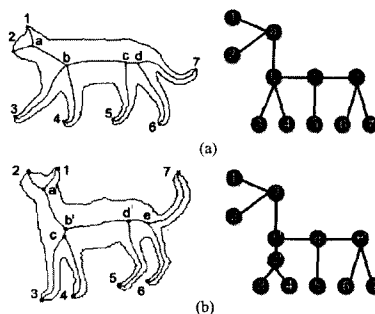


图1 同类物体的骨架图

在描述算法前,先定义本文将要用到的3个概念:(1)端节点:若一个骨架点在骨架图上只存在一个点与其相邻,则该骨架点被称为一个端点;(2)接合节点:若一个骨架点存在3个或更多的相邻点,则称其为接合节点;(3)骨架路径:骨架上某个接合节点到骨架某个端节点的最短路径,称为骨架路径。

本文采用骨架的接合节点作为骨架路径的起始点,主要原因是:当使用接合节点做骨架路径的起点时,距离重心的信息将会增多;而端节点做骨架路径的起点时,骨架分支的信息将会增多,我们容易观察到骨架分支在很多情况下是不稳定

到稿日期:2010-02-05 返修日期:2010-04-21 本文受国家自然科学基金(60903096,60873127),国家科技支撑计划(2006BAH02A24)等项目资助。

陈 珺(1980—),女,博士生,讲师,主要研究方向为计算机视觉与机器学习,E-mail:chenjun71983@163.com;王 波(1987—),男,本科,主要研究方向为计算机视觉;周 瑜(1983—),男,硕士,主要研究方向为机器学习;白 翔(1981—),男,博士,主要研究方向为计算机视觉;刘文予(1963—),教授,博士生导师,主要研究方向为多媒体通信与多媒体信息处理、媒体内容与安全、视觉计算与计算机图形。

的,如图2所示的两个骆驼图片。图片蓝色显示的区域为骆驼的躯干部分,骨架的接合节点主要在这个区域;图片的红色区域主要是骨架端节点所在区域。从图2可以看到,红色区域主要包含骆驼的头部、腿部、驼峰和尾巴,两幅图中,这些地方变化是很大的,也就是说这些部位的信息在匹配中占有的比例较大时,会使匹配变得不稳定,所以应当减少从这些部分提取的信息在匹配中所占有的比例。因此骨架接合节点的相似性更能反映出形状的相似性。

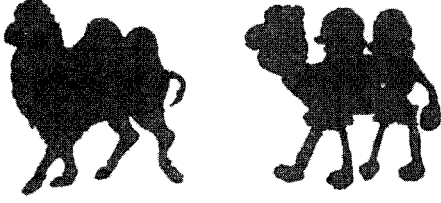


图2 躯干部分骨架信息更稳定示意图

2.1 骨架路径的距离的计算

由上述可知,对于每一个骨架接合节点,都可以找到其与每个骨架端节点的骨架路径^[5]。因此节点的相似性求解就转化为骨架路径间的距离求解。我们这里主要研究接合节点到端节点的骨架路径。用 sp 代表骨架路径,在 sp 上取 M 个采样点, l 代表路径长度, R 表示 M 个骨架路径采样点上的最大盘半径:

$$R = (r_1, r_2, \dots, r_M) \quad (1)$$

这里半径 $R(S)$ 近似于每个骨架 S 上的距离变化 $EDT(s)$ 。假设原始形状上有 N_0 个像素点,为了让该方法具有尺度不变性,通过以下方法对 $R(S)$ 进行归一化:

$$R(S) = \frac{EDT(s)}{\frac{1}{N_0} \sum_{i=1}^{N_0} EDT(s_i)} \quad (2)$$

式中, $s_i, i=1, 2, \dots, N_0$ 代表了原始形状上全部 N_0 个像素点。两段路径的形状相似性被称为路径距离。 R 和 R' 分别表示两段路径上的半径向量, l 和 l' 分别表示两段路径的长度,同样也是归一化以后得到的,这里是将所有路径长度对最长的一条路径进行归一化。最终可以得到路径 sp 的距离 pd :

$$pd(sp, sp') = \sum_{i=1}^M \frac{(r_i - r'_i)^2}{r_i + r'_i} + \beta \frac{(l - l')^2}{(l + l')} \quad (3)$$

式中, β 为权重因数。

2.2 骨架的接合节点匹配

对图1所示的两幅骨架图,其接合节点序列分别为 (a, b, c, d) 和 (a', b', c', d', e') ,接合节点序列个数不相同,为匹配带来了困难。

我们首先用最佳序列双射方法^[4]进行匹配。该方法旨在弹性匹配两个长度不同的序列。也就是说,它可以略去匹配序列中的一些异常元素,在这里意味着匹配时略去骨架的一些接合节点。假设两个序列的长度分别为 m 和 n , $a = (a_1, a_2, \dots, a_m)$ 和 $b = (b_1, b_2, \dots, b_n)$ 。最佳序列双射算法可以找到 a 中的子序列 a' 和 b 中的子序列 b' 的最佳匹配关系。这里定义一个对应关系函数 $f: \{1, \dots, m\} \rightarrow \{1, \dots, n, \infty\}$,它是有限域函数 $f: \{1, \dots, m\} \rightarrow \{1, \dots, n\}$ 的一个单调双射。例如,一个函数 f 满足 $f(i) < f(i+1) < \infty$,对于所有 $i \in \{1, \dots, m\}$,有 a_i 被映射到 $b_{f(i)}$,并且对于映射到 ∞ 的元素允许不只一个。 $f(i) = \infty$ 表示序列 a 中的第 i 个元素被跳过。系数集合

(i_k) 和 $(f(i_k))$ 满足对所有 $i_k \in \{1, \dots, m\}$ 有 $f(i_k) < \infty$ 成立,就这样定义了 a 的子序列 a' 和 b 的子序列 b' ,使得受限子 (i_k) 的 f 是一个双射。以上解释了最优子序列双射中“子序列双射”的含义。然而,仍需给出“最优”的定义。设距离函数 d 表示任意两个元素 a 和 b 之间的非相似度,例如 $d(a_i, b_j)$ 表示 a 中第 i 个元素到 b 中第 j 个元素之间的距离,这里 $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n, \infty\}$ 。对距离函数 d 并没有特定的限制,因此任意的距离函数 d 都是适合的。这里,距离函数 d 就是式(3)中的路径距离 pd 。当 $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$ 时,可直接获得 $d(a_i, b_j)$,而 $d(a_i, \infty)$ 的取值是十分关键的。对于所有 $i \in \{1, \dots, m\}$, $d(a_i, \infty)$ 通常是一个常数,它决定了跳过序列 a 中的元素之后的耗费大小。这里称其为跳跃值($jumpcost$)。

$$jumpcost = mean_i (\min_j (d(a_i, b_j)) + std_j (\min_j (d(a_i, b_j)))) \quad (4)$$

因此,对于每一个元素 a_i ,找到其距离最近的 b_j ,然后用所有对最近的元素间的距离平均值加上距离的标准差(std)来获得跳跃值。例如,若在序列 a 和 b 中,除了一个异常元素 a_k 之外其它元素十分相似,那么每一个 $a_i (i \neq k)$ 到一个元素 b_j 都有一个很小的距离。由此可知跳跃值也很小,使得 a_k 到 a 中元素的最近距离将会大于跳跃值,那么元素 a_k 将会被排除在匹配范围之外且带来一个相对较小的惩罚,正如匹配时跳过该元素。对于任意给定的对应关系,可以定义两序列之间的距离为:

$$d(a, b, f) = \frac{1}{m} \sum_{i=1}^m d(a_i, b_{f(i)})^2 \quad (5)$$

OSB的目标是要找到一个对应关系 f ,使得 $d(a', b', f)$ 的值最小。更加准确地说,要找到一个最优的对应关系 f ,它的 $d(a, b, f)$ 在所有可能的对应关系 f 中是全局最小的:

$$\hat{f} = \operatorname{argmin} \{d(a, b, f) : f \text{ is a correspondence}\} \quad (6)$$

最优的匹配对应关系可以通过在一个有向图上用最短路径算法来寻找。

经过最佳序列双射法可以得到图1中接合节点匹配子序列 (a, b, c, d) 和 (a', b', d', e') 。以接合节点为起始点,对图1(a)所示的骨架图,可以得到对应4个接合节点的所有骨架路径:

$$A = \{a1, a2, a3, a4, a5, a6, a7\}$$

$$B = \{b1, b2, b3, b4, b5, b6, b7\}$$

$$C = \{c1, c2, c3, c4, c5, c6, c7\}$$

$$D = \{d1, d2, d3, d4, d5, d6, d7\}$$

其中, A, B, C, D 分别表示图1(a)中骨架的4个接合节点 a, b, c, d 对应的骨架路径集。

对图1(b)的骨架,以接合节点为起始点,也得到4条骨架路径集:

$$A' = \{a'1, a'2, a'3, a'4, a'5, a'6, a'7\}$$

$$B' = \{b'1, b'2, b'3, b'4, b'5, b'6, b'7\}$$

$$D' = \{d'1, d'2, d'3, d'4, d'5, d'6, d'7\}$$

$$E' = \{e'1, e'2, e'3, e'4, e'5, e'6, e'7\}$$

2.3 骨架的相似性计算

对图1所示两幅骨架图,由式(3)可以计算任意两条骨架路径的距离,进而得到一个路径距离矩阵。例如,接合节点 a 和 a' 的距离矩阵为:

$$pd(A, A') = \begin{bmatrix} pd(a_1, a'_1) & pd(a_1, a'_2) & \cdots & pd(a_1, a'_7) \\ pd(a_2, a'_1) & pd(a_2, a'_2) & \cdots & pd(a_2, a'_7) \\ \vdots & \vdots & \ddots & \vdots \\ pd(a_7, a'_1) & pd(a_7, a'_2) & \cdots & pd(a_7, a'_7) \end{bmatrix} \quad (7)$$

再利用本文上一节中的最佳序列双射方法,可以得到两个接合节点的非相似度:

$$c(A, A') = OSB(pd(A, A')) \quad (8)$$

这个非相似度反映了骨架节点的不相似性。即经过计算,可以得到两个骨架所有接合节点的距离矩阵:

$$DIS = \begin{bmatrix} AA' & AB' & AD' & AE' \\ BA' & BB' & BD' & BE' \\ CA' & CB' & CD' & CE' \\ DA' & DB' & DD' & DE' \end{bmatrix} \quad (9)$$

如果两幅图像相似,其接合节点在图中的位置也相似,则它们之间的距离也越小;反之,它们的距离就很大。

3 训练神经网络

假设某个数据库一共有 N 个类,每个类有 T 个形状,共有 $N \times T$ 个形状。每类中取一个形状做测试,其它的 $W = N \times (T-1)$ 个形状做训练样本。训练时,取每类中节点数最多的一个形状作为匹配基,第 i 类中的匹配基的接合节点数记为 J_i , N 个类共有 s 个基节点: $s = \sum_{i=1}^N J_i$ 。第一个形状中与第一个基节点 J_1 的距离最近的接合节点为 K_{1X} ,它们之间的距离用 $K_{1X}J_1$ 表示,这里 X 表示距离最小的未知接合节点。因此,对于每一个形状都可以得到一个 s 维的向量,如图 3 所示。

$$\begin{aligned} A_1 &= \{K_{1X}J_1, K_{1X}J_2, \dots, K_{1X}J_s\} \\ A_2 &= \{K_{2X}J_1, K_{2X}J_2, \dots, K_{2X}J_s\} \\ &\vdots \\ A_w &= \{K_{wX}J_1, K_{wX}J_2, \dots, K_{wX}J_s\} \end{aligned}$$

图 3 神经网络的输入向量基

采用 BP 神经网络模型,将训练样本送入神经网络的输入端,并在样本输出端输入样本所属的类别,训练过程如图 4 所示。

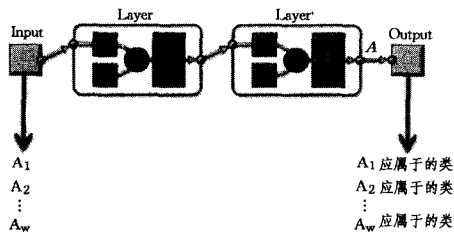


图 4 训练用于分类的神经网络

4 实验结果

为了验证匹配识别的效果,在标准数据库 Tari^[1]的两个数据库上进行测试。第一个数据库共有 56 个形状、14 个类,每类有 4 个形状。第二个数据库中共有 180 个形状、15 个类,每类 12 个形状。在对第一个数据库进行实验时,每类中选一个形状,共 14 个形状做测试。在 42 个测试样本中,每类选取节点最多的形状做匹配基,组成向量空间。这里使用参

数 $M=50, \beta=40$, 训练速率=100, 训练目标=0.0001。表 1 是本文实验结果与 IDSC+DP^[6] 算法实验结果的对比,可以明显看出本文的算法优于 IDSC+DP 算法,并且具有更好的稳定性。图 5 示出对实验数据库进行 4 次训练后均方差的变化情况。可以看到,均方差下降很快,达到了理想的效果。

表 1 Tari56 数据集的分类正确数结果比较

算法	第一次 实验正确数	第二次 实验正确数	第三次 实验正确数
IDSC+DP	53	51	38
本文方法	54	54	54

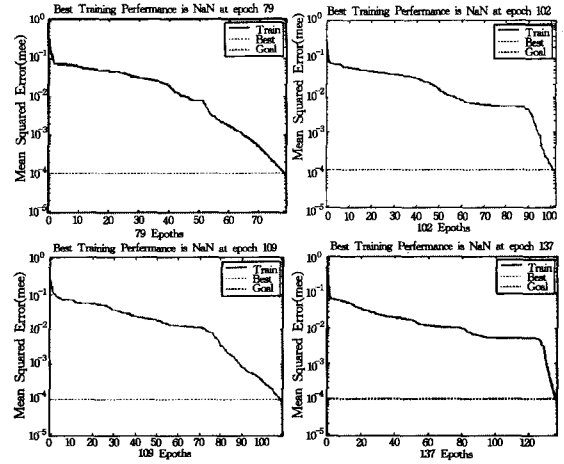


图 5 Tari56 数据集的 4 次分类训练的均方差

在对第二个数据库数据进行实验时,同样每类选取一个形状做测试,其它的做训练,这里使用参数 $M=50, \beta=40$, 训练速率为 100, 训练目标为 0.0001。分类正确率达到 100%。可以看到,当训练样本越多时,分类正确率越高。

结束语 本文提出了一种新的形状分类方法,它将形状与形状的距离转化为多维的形状骨架节点间的距离向量,使其比一维的距离有更强的易用性。通过序列双向匹配算法,排除了多余的骨架分支,使得算法具有更强的鲁棒性。并且利用神经网络对高维向量的处理能力提出了一个新颖的形状分类算法。该方法在两个标准数据库上进行的测试,均取得了较好的效果。而且,当实验的学习样本较多时,该方法体现了非常强大的分类性能。

参考文献

- [1] Aslan C, Tari S. An axis based representation for recognition [C]//Proc Int'l Conf. Computer Vision. 2005;1339-1346
- [2] Bai X, Latecki L J, Liu W Y. Skeleton pruning by contour partitioning with discrete curve evolution [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2007, 29: 449-462
- [3] Bai X, Latecki L J. Path similarity skeleton graph matching [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(7): 1282-1292
- [4] Latecki L J, Wang Q, Koknar-Tezel S, et al. Optimal Subsequence Bijection [C]. IEEE Int. Conf. on Data Mining. 2007
- [5] Xu Y, Wang B, Liu W Y, et al. Skeleton Graph Matching Based on Critical Points Using Path Similarity [C]//Asia Conference on Computer Vision (ACCV). 2009
- [6] Ling H B, Jacobs D W. Shape classification using inner-distance [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29: 286-299