

链码和在边界形状分析中的应用

陆宗骥 童 韬

(华东理工大学信息学院, 上海 200237)

摘要 介绍了应用链码分析边界形状的新方法. 通过引入相对链码与绝对链码概念得到了计算链码和(平均链码)的简便算法, 边界上连续3点的绝对链码之和可以表示边界点的切线方向(斜率), 进入和离开边界点的3点链码和之差可以表示边界的曲率. 同时给出了利用这些参数判别边界角点、边界光滑段的判据, 以及估算曲率半径、等效周长的方法; 最后, 以细胞边界凹陷的修补和重叠细胞粘连的分割为例介绍了算法具体的使用, 该算法在3类细胞中采集到的20余组粘连和缺损细胞上进行了验证, 结果表明, 该算法处理速度快, 分割效果良好.

关键词 链码 微分链码 链码和 形状分析

中图法分类号: TP391.4 **文章标识码:** A **文章编号:** 1006-8961(2002)12-1323-06

The Application of Chain Code Sum in the Edge form Analysis

LU Zong-qi, TONG Tao

(College of Information, East China University of Science and Technology, Shanghai 200237)

Abstract This dissertation expands on a new algorithm, which has the function of analyzing edge form with the chain code. By the introduction of the concepts of relative chain code and absolute chain code, we proposed a simple and direct algorithm to compute chain code sum (average chain code). We found that tangent direction (slope) of edge point can be figured by the absolute chain code sum of three sequential points and curvature of edge can be figured by the difference of chain code sum between the three sequential points ingoing and outgoing. Furthermore, we provided the criterion distinguishing edge inflexion and sleek curve section and the method computing inaccurately curvature radius and approximate perimeter. In the end of the paper, we introduced the use method with cell edge hollow repairing and overlap or conglutination cell segmenting for example. This algorithm has proved high-speed and has a good effect of cell segmentation on more than twenty groups of conglutinate and absent cells which are gathered from three kinds of cells.

Keywords Chain code, Differential chain code, Chain code sum, Form analysis

0 前言

在医学图象处理中, 细胞图象的处理与分析也是应用重点之一. 例如, 对于神经细胞图象, 通过细胞形状的分析可区分细胞的类型; 对于免疫组化细胞图象, 根据区域内阴、阳性细胞数的比率可判定病人的患病程度; 对于染色细胞 DNA 含量分析, 细胞光密度的累加和与其 DNA 含量成正比, 等等. 而在这些处理中, 第 1 步必须先 from 图象中把细胞分割出来. 由于细胞图象通常比较复杂, 分割后得到的细胞

轮廓往往存在粘连和缺损, 这就给以后的分析造成了很大的困难, 遇到这种情况, 通常需要人工进行干预, 即先将细胞粘连处划线分开, 缺损处则需连线补全. 这在实际使用时极不方便. 在粘连情况严重时, 由于所需干预太多, 甚至使处理成为不可能. 图 1 给出了若干细胞粘连和轮廓缺损的例子. 从图 1 可以看出, 自动分割时, 理想的分割点大都处于轮廓的凹、凸角点处, 自动分割的关键是能否迅速地找出这些角点.

文献[1]给出了一个采用链码分析多边形形状的方法. 多边形的每个顶点在微分链码曲线上对应

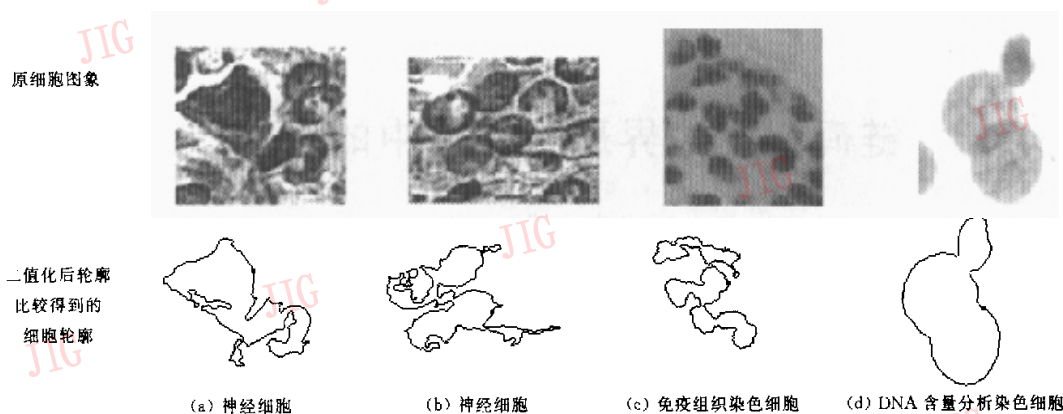


图1 细胞图象及细胞轮廓

一个尖峰凸起,正峰处对应凸边,负峰处对应凹边,从而可从微分链码曲线图中找出多边形顶点的位置.链码把整个圆周只分成8个等份,分度太粗,对其文中只介绍了原理,并未介绍具体应用.

文献[2]用边界序列中相隔数点连线的方向表示边界的斜率,并进而计算该处的曲率.设 S 是区域的边界,它由 n 个点组成,即

$$S = p_1 p_2 \cdots p_i \cdots p_n$$

S 在 p_i 点的左向(或右向) k 步斜率是指从 p_i 到 p_{i-k} (或 p_{i+k})的方向; S 在 p_i 点的 k 步曲率是指 p_i 的左、右向 k 步斜率之差.曲率的局部极值点称为 S 的角点,边界的角点在形状分析中是一个重要的位置特征.通常, k 取值为3,即常采用3步斜率与3步曲率.这种方法十分有效,但比较注重曲率与角点的作用,而忽略了斜率的作用,同时计算方法也略为复杂了些.

文献[3]介绍了利用平均链码表示角度的新方法,当取平均的像素数固定时,链码和也可表示角度,通常求和的像素数取3.可用边界上连续3点的链码之和表示该边界点的切线方向(斜率),可用进入与离开该边界点的3点链码之和之差表示边界的曲率.

本文改进了文献[3]计算链码和(平均链码)的算法.并得出利用这些参数判别边界角点、边界光滑段的判据,以及估算曲率半径、等效周长的方法,进而实现粘连细胞的自动分割和缺损轮廓的自动补全.

1 三点链码和及其计算方法

1.1 链码、平均链码、链码和

链码的定义见图2,它用中心像素指向它的8

个邻点的方向来定义的.链码值加一,其所指方向就逆时针旋转 45° ,链码值递增,其所指方向就按逆时针方向旋转.因为每个像素只有8个邻点,所以链码值只有 $0 \sim 7$ 8个^[4].

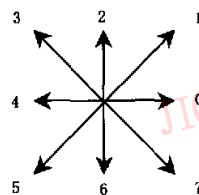


图2 链码的定义

链码序列的平均值为平均链码,直线的平均链码可以表示它的方向.平均链码的取值范围为 $0 \sim 8$ (不包括8).在定义链码的8个方向上,直线的平均链码取整数,其他方向上的直线的平均链码则取小数.为了准确地表示方向,在计算链码值 $1, 0, 7$ 方向附近直线的平均链码时,需作一些附加的规定,即相邻方向的链码应取相邻整数值.当计算平均链码的点数固定时,可以不用取平均,这时链码和就可用来表示直线的方向^[3].由于边界的一个局部可以近似看作直线,所以边界上连续的几个链码之和可以近似地表示该处的斜率.

为了得到简便的计算方法,特引入相对链码、绝对链码两个新概念.

1.2 相对链码、绝对链码

对于边界上的每一个点都有前一点指向此点的链码 C_1 和此点指向后一点的链码 C_2 ,所谓相对链码是指 C_2 和 C_1 的相互关系,即当它们方向一致时,它们的相对链码为0.如果 C_2 相对于 C_1 作逆时针旋

转时, 相对链码按偏角的大小取值 1~4, 分别对应逆时针旋转 45°、90°、135°和 180°; 作顺时针旋转时, 相对链码取值 -1~-3, 分别对应顺时针旋转 45°、90°和 135°。

所谓绝对链码是指从起始点开始相对链码的累加值, 起始点的绝对链码设为 0。这样, 沿着边界移动一周回到起点, 其绝对链码值增加 8。引入了绝对链码概念后, 求链码和时就不用再考虑 0 和 7 交界处链码的取值问题了。

1.3 绝对链码及三点链码和的计算

设 $C(i)$ 和 $C(i-1)$ 分别为当前点 i 与前一点 $i-1$ 的链码, $R(i)$ 为它们的相对链码, $A(i)$ 和 $A(i-1)$ 分别为当前点和前一点的绝对链码。则绝对链码的计算过程如下:

$$A(0) = 0 \quad (1)$$

$$R(i) = [C(i) - C(i-1) + 8] \text{ MOD } 8 \quad (2)$$

$$\text{IF } (R(i) > 4) \text{ THEN } R(i) = R(i) - 8 \quad (3)$$

$$A(i) = A(i-1) + R(i) \quad (4)$$

当边界的点数为 N 时, 边界点的序号分别为 $0 \sim N-1$ 。从起始点开始, 逆时针绕边界走一圈回到起始点时其绝对链码之差为 8, 即

$$A(N) - A(0) = 8 \quad (5)$$

三点链码和即为当前点和前两点的绝对链码之和, 即

$$\text{Sum}(i) = A(i) + A(i-1) + A(i-2) \quad (6)$$

由于边界是闭合的, 起始两点计算时, 前面点的值要绕到尾部去取, 此时需注意绝对链码的调整值 8, 即

$$\text{Sum}(0) = A(0) + A(N-1) + A(N-2) - 16 \quad (7)$$

$$\text{Sum}(1) = A(1) + A(0) + A(N-1) - 8 \quad (8)$$

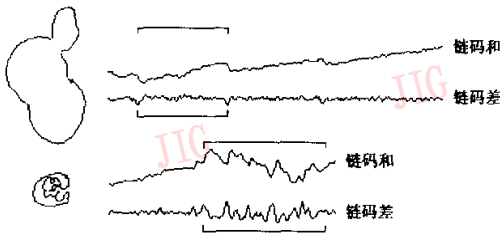


图 3 细胞边界及其链码和曲线与链码差曲线

如此计算得到的 $\text{Sum}(0)$ 未必等于 0。为了统一起见, 便于比较, 再将所有的 $\text{Sum}(i)$ 值都减去原 $\text{Sum}(0)$ 值, 由此即得以 0 开始的三点链码和序列, 见图 3 右边两组曲线中的上面一条。其中, 边界上小

方块的中心点为边界起点, 按逆时针方向进行跟踪, 上下水平标志线的间隔为 24。三点链码和将圆周分为 24 等份, 对于边界的光滑部分可以表示切线的方向, 其值相差 1, 相当于相差 15°。类似的有, 从起始点开始, 逆时针绕边界走一圈回到起始点时, 其三点链码和之差为 24, 即

$$\text{Sum}(N) - \text{Sum}(0) = 24 \quad (9)$$

为简便起见, 三点链码和以后简称链码和。

2 边界的基本参数

在边界的形状分析中需要用到不同的特征参数, 如边界点的坐标、斜率、曲率与边界的累积长度等等, 它们都可由边界的链码计算得到, 通常, 在形状分析之前预先将这些参数计算出来, 以供分析时随时取用。

(1) 链码

边界跟踪的结果可得到它的链码。通常, 链码的数据结构为: 第 1、2 个单元为起始点的 x, y 坐标, 第 3 单元为边界点数 N , 后面单元依次存放各点的链码^[1]。这里, 边界跟踪取逆时针方向。

(2) 坐标

从起始点开始, 由边界的链码可依次计算得到各点的坐标, 由此可把边界绘制出来。链码序列表示了边界上近邻点间的邻接关系, 但无法表示序列上其他点在位置上的相互关系。利用坐标就可计算得到边界上任意两点间的距离(弦长), 从而判别它们的相对位置。这在曲线段的连接和配对时尤为重要。

(3) 链码和

链码和就是没有归一化的平均链码, 它可表示边界的方向(斜率), 即当前点的切线方向, 见图 4 中的虚线所示方向, 它要比原来的链码精确些, 其中前者将圆周分成 24 等份, 后者仅分成 8 等份。因此, 与文献[1]相比, 用链码和值连成的曲线可以更好地反映边界的形状特征。链码和与文献[2]中的 3 步斜率

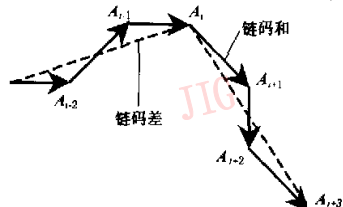


图 4 i 点处的链码和与链码差

相当,但它的计算更为简便.

(4) 链码差

链码差是指离开与进入此点的三点链码和之差(如图 4 中两虚线所示方向之差),即

$$Diff(i) = Sum(i + 3) - Sum(i) \quad (10)$$

同样,末尾 3 点计算时,数据应从起始点处取,并且需作相应调整,即

$$Diff(N-i) = Sum(N-i+3) - Sum(N-i) \\ = Sum(3-i) + 24 - Sum(N-i) \quad (11)$$

其中, $i=1, 2, 3$.

链码差表示两个方向间的差值,是一个与曲率成正比的量,可用于发现边界上的角点.当边界走向为逆时针方向时,差值为正的点为凸角点,差值为负的点为凹角点.链码差与文献[2]中的 3 步曲率相当.以图 4 中的链码为例计算它们的链码和与链码差,其结果见表 1.其中,假定 A_{-2} 为 0. Sum_i 表示第 $i-3$ 点至 i 点的方向, Sum_{i+3} 则表示第 i 点至 $i+3$ 点的方向.两者之差 $Diff_i$ 则表示第 i 点处方向的变化,其值为 -5 说明轮廓方向顺时针旋转了 75° .

表 1 链码和与链码差的计算

.....	A_{-2}	A_{-1}	A_i	A_{i+1}	A_{i+2}	A_{i+3}	Sum_i	Sum_{i+1}	Sum_{i+2}	Sum_{i+3}	$Diff_i$
	0	1	0	-1	-2	-1	1	0	-3	-4	5

图 3 给出了两个细胞的边界,以及它们的链码和曲线与链码差曲线.链码和表示边界的切线方向;链码差表示边界的曲率.上方大细胞的曲线,图框部分表示右上方重叠细胞的链码段,它的起点是两个凹角点.下方细胞的曲线,图框部分表示细胞上凹陷段的链码,它们的起点对应两个凸角点.从图 3 可以看出,与文献[1]相比较,它能较为精确地反映边界的方向和曲率信息.

(5) 累积长度

累积长度是指边界起点到当前点之间的边界长度,边界上任意两点间的弧长等于它们各自的累积长度之差.为了计算方便起见,水平、垂直向的步长设为 1,斜角方向上的步长设为 1.5.累积长度可在形状分析时用于计算某段边界的弧弦比.边界的弧弦比可以用来判别该边界段是否属于另一物体,或者判别它是否为原边界中的一个凹陷.

3 形状特征

(1) 链码和曲线的特征

在链码和曲线上,水平向的直线段对应边界上

的直线,倾斜直线对应边界上的圆弧,倾角小的线段对应边界上半径大的圆弧.倾角向上时,为沿逆时针方向转;倾角向下时,为沿顺时针方向转.圆弧两端点的链码和之差相当于此弧段所对的圆心角.图 3 上图重叠细胞的链码和曲线反映出它有 3 段大的圆弧边界.其中,第 1 弧段所对的圆心角较大,故其等效直径较小.

(2) 角点、边界光滑部分和边界段所对圆心角的估算

判定角点的判据为

$$Diff(i) < -1 \text{ 或 } Diff(i) > 2 \quad (12)$$

凹角点处 $Diff$ 值为负,其值 -2 相当于方向变化 30° ;凸角点处 $Diff$ 值为正,其值 3 相当于方向变化 45° .因为轮廓粘连处多数呈现凹角点,缺损处多数呈现凸角点,而在细胞图象中出现轮廓粘连的情况远比出现缺损的情况为多,为防遗漏,判别条件放宽些.

判定边界光滑段的判据为

$$(m-n) > 15 \text{ AND } [Sum(m) - Sum(n)] > 3 \text{ AND } -2 < Diff(i) < 6 \quad (n \leq i \leq m) \quad (13)$$

其中, m, n 为光滑段端点的序号.

图 5(a)为原始边界图,图 5(b)中的黑线为边界的光滑部分,黑点为边界上的角点.

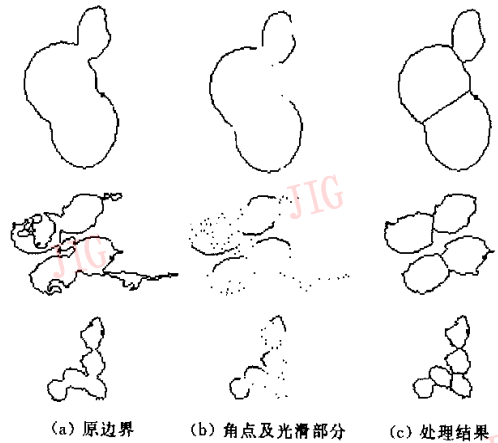


图 5 细胞原始边界图,光滑部分及处理结果

边界段所对圆心角的估算公式为

$$A = 15[Sum(m) - Sum(n)] \quad (14)$$

(3) 边界点曲率半径和等效周长的估算

边界光滑段曲率半径的估算公式为

$$R = \frac{12}{\pi} \frac{L(m) - L(n)}{Sum(m) - Sum(n)} \quad (15)$$

对于圆周来说,任意两点的累加长度之差等于它们之间的弧长,弧两端点处切线的夹角等于此弧所对的圆心角,而由弧长及其所对的圆心角可计算出圆的半径.式(15)则是利用这个关系由轮廓的光滑段来估算轮廓的等效半径,其中, $L(m)$ 、 $Sum(m)$ 、 $L(n)$ 和 $Sum(n)$ 分别为光滑部分两个端点处的累积长度与链码和,因子 $12/\pi$ 也可用 4 代替.图 5 给出 3 组重叠、粘连及边界缺损细胞,此图中的 3 类细胞用式(15)式计算得到的平均曲率半径分别为 30、18 和 11.类似地,等效圆周长(S)可用下式估算

$$S = 24 \frac{L(m) - L(n)}{Sum(m) - Sum(n)} \quad (16)$$

曲率半径和等效圆周长估算的精度不是很高,但在区分不同类型细胞时用上基本上可将它们的大小彼此区分开来.

(4) 弧弦比

边界上两点之间的长度与它们间的距离之比称为弧弦比.在确定粘连细胞的分割位置时,可用它作判据,如图 5(a)上面细胞右上方粘连部分的分离,最佳位置应在两个凹角点附近.此时除了它们的曲率较大外,弧弦比也应为局部极大值.

4 边界分析实例

以医学图象处理中,细胞边界凹陷的修补以及粘连细胞的分割为例说明上述特征参数的实际应用.处理步骤如下:

(1) 确定细胞类型

根据边界光滑段估算细胞的等效半径(或等效周长),由此确定细胞的类型.本文涉及的 3 种细胞间分隔值为 23 和 13,即第 1 类细胞的等效半径大于 23,第 3 类细胞的等效半径小于 13,第 2 类细胞的等效半径介于 13~23 之间.

(2) 细胞边界凹陷处的修补

边界凹陷段端点间的距离一般比较小,此处取小于 6,同时它们通常是链码差取正值的角点.这样,在满足上述条件的角点间两两计算弧弦比,在弧弦比取局部极值,并且数值大于 4 的地方,凹陷应该修补,见图 6 中的第 1 步,图中同时给出了相应边界的链码和曲线与链码差曲线,曲线图上加框处指出了修补的位置.

(3) 细胞小粘连的分割

所谓小粘连是指弦长上限约为细胞周长的

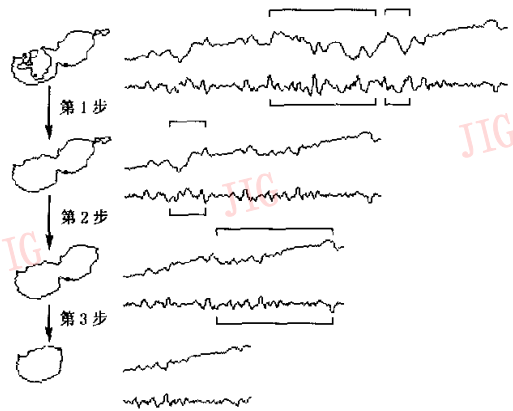


图 6 细胞边界凹陷的修补和粘连的分割

$\frac{1}{15} \sim \frac{1}{20}$,对于图 5 所示 3 类细胞分别取 20 个、8 个和 3 个像素,而且边界段中至多只能包含 1 个细胞.边界的粘连部位通常在取负值的角点处,它们的弧弦比极值可取在 2.5~4 之间,对于大的细胞,此比值可取得小些.细胞小粘连的分割见图 6 中的第 2 步,以及图 7 中的第 1 步.

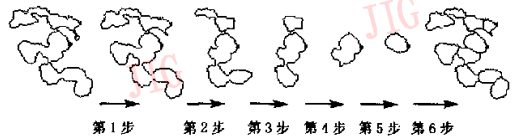


图 7 多细胞粘连的分割过程及处理结果

对于大而边界形状复杂的细胞,以上两类处理仅采用上述条件还不够,还需在修改时考虑前、后连接部分在方向上的一致性,以免造成误判.

(4) 大粘连与多细胞粘连的分割

所谓大粘连是指弦长上限约为细胞周长的 $\frac{1}{3} \sim \frac{1}{5}$,对于图 5 所示 3 类细胞分别为 50 个、20 个和 15 个像素.弧弦比也可取在 2.5~4 之间.同时,除了弦长可能稍长外,大粘连又指边界段中可能包含多个细胞.由于需要考虑多细胞粘连,分割应是一个串行处理过程.因需要判别构成细胞的条件,只能由底层开始逐次分步进行分割,即先行分割的细胞中不能嵌套别的细胞(见图 7).

(5) 细胞的判别

分割结果根据其大小与形状条件确定是杂质还是细胞.图 6、图 7 显示了细胞边界修补、分割处理的过程及其处理结果.

5 结 论

本文介绍了三点链码和的几何意义与计算方法,以及它和它的派生参数在边界形状分析中的应用,并以细胞边界的修补和分割为例作了说明.本算法在3类细胞中采集到的20余个细胞边界上进行了验证,效果良好,并将处理过程制作了一个演示程序.今后,有待于应用到实际处理软件中去,并加以完善,它也有望应用于其他需要修补边界凹陷、分割边界粘连的场合.

参 考 文 献

- 1 Castleman K R. Digital image processing[M]. 北京:清华大学出版社,1998;479~480,496~497.
- 2 徐建华. 图象处理与分析[M]. 北京:科学出版社,1992;176~177.
- 3 陆宗祺等. 工程图纸矢量化中平均链码与线条轮廓跟踪[J]. 模式识别与人工智能,1997,10(2):147~152.
- 4 Freeman H. Computer Processing of Line-drawing Image[J]. Computing Surveys, 1974,6(1):57~97.
- 5 冈萨雷斯 R C. 数字图象处理[M]. 北京:科学出版社,1982.



陆宗祺 1945年生,华东理工大学信息学院教授,研究方向为图象处理与模式识别.主要致力于计算机图象知识的普及及图象处理传统算法的研究.编有DOS、Win32,以及C++ Builder环境下的图象处理演示程序.



董 轲 1977年生,华东理工大学信息学院硕士研究生,研究方向为数字图象处理和多媒体技术在工业检测中的应用.