

# 基于基因搜索的模糊图象解释方法

钱 涛 王 琦

(浙江大学计算机科学与工程系, 杭州 310028)

**摘 要** 图象解释是计算机视觉的重要组成部分, 它涉及图象处理、分类器设计和逻辑推理等许多领域. 针对目前图象解释系统要面对的严重噪声、模糊性和不确定性问题, 重点研究了一种基于基因搜索的双向推理技术, 该算法分为如下两步: 首先通过基于分割区域统计/几何特征的模式分类器来得到初始的分类模糊隶属度, 并根据经验(或统计)得到的先验空间位置关系模糊规则来构造一种有效表达图象解释信息的模糊图; 然后通过基因搜索算法融合上面的两类信息来得到图象的最佳解释. 实验结果表明, 该方法对具有单一对象或多个对象的区域均有很好的效果, 也是对基于概率、证据和模糊推理等单向推理机制图象解释方法的提高.

**关键词** 模糊图 基因搜索 图象解释 空间位置关系

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2002)03-0218-05

## Fuzzy Image Interpretation Based on Genetic Searching

QIAN Yun-tao, WANG Qi

(Department of Computer Science and Engineering, Zhejiang University, Hangzhou 310028)

**Abstract** Image interpretation is an important part of computer vision, which is related to many fields such as image processing, classifier designing and logic reasoning. In this paper, genetic searching based two-directional reasoning is discussed. The algorithm consists of two steps. At first, the fuzzy memberships of classification is obtained by fuzzy classifier based on the statistic/geometric features of segmented regions, and a fuzzy graph used for effectively representing interpretation information is constructed through prior rule-base concerning about spatial relations that is acquired from statistics or experience. At second, genetic searching algorithm is used to combine the above two types of information, and the optimistic interpretation is achieved. In order to decrease the computational cost and increase the possibility of getting optimal solution, a new crossover operator of genetic searching is proposed that is based on non-random graph partition. The experiments show that genetic searching based fuzzy image interpretation is powerful for the regions that include one or more objects. This method is an improvement over the one-directional reasoning method based image interpretation such as probabilistic, evidence and fuzzy reasonings.

**Keywords** Fuzzy graph, Genetic searching, Image interpretation, Spatial relation

## 0 引 言

图象解释是计算机视觉中最富有挑战性的研究方向, 在航空/卫星照片分析、自动导航系统和智能诊断等领域具有广泛用途<sup>[1~3]</sup>. 按照采用的信息类型, 图象解释方法可分为如下两类: (1)单纯基于对

象统计/几何特征的方法; (2)基于统计/几何特征和对象之间关系特征(主要是空间位置关系, 这也是本文讨论的重点)的方法. 其中, 第 1 种方法对对象特征的精度和可靠性均有较高要求, 而在大多数图象解释系统中, 由于环境和图象摄取设备的限制, 以及图象处理和分割算法的非绝对可靠性, 这一条件很难满足; 第 2 种方法由于融入了对象之间空间关系

约束信息, 因此解释具有更好的鲁棒性和准确性, 已被广泛采用。

上述图象解释的困难在于, 经过一系列处理得到的对象特征不是完全正确的, 有些对象本身并没有明确的定义或者对象之间的特征界线不明显, 这样, 基于统计/几何特征的模式分类结果就包含了很严重的噪声、模糊性和不可靠性。另外, 由于对象关系特征的先验知识也是不完全、不清晰和不绝对的, 因此目前图象解释系统广泛采用 Bayes 网络、证据推理网络和模糊神经网络等不确定性推理方法<sup>[4,5]</sup>, 这些网络都是有向无圈图结构, 节点由假设空间和证据空间构成, 虽假设节点可以转换成另一些假设的证据, 但不能构成回路; 而在基于对象相互空间位置关系的不确定推理中, 则是假设空间和证据空间是不可分割的, 这样推理网络就是有向回路结构。由于现在这种对象之间互相影响的解释系统缺乏良好的数学工具, 因此本文提出了一种基于基因搜索和模糊空间位置关系规则的图象解释方法。

## 1 模糊图结构

模糊图适合表达对象之间复杂的关系, 它由节点集  $V$  和边集  $E$  构成, 即  $G=(V, E)$ , 假设对象空间为  $X$ , 则图中每个节点表示对象空间上的一个模糊子集, 而在图象中, 则代表一个分割区域, 该模糊子集可由基于统计/几何特征的模式分类器得到。另外, 在节点集中添加一个特殊节点  $x_0$ , 它代表整幅图象。

边分为如下两类:

第 1 类边是无方向的, 它表示节点所代表的区域之间的空间位置关系,  $v_i$  和  $v_j$  ( $0 < i < j$ ) 是任意两节点, 且它们之间的边由 4 个元素构成, 即

$$e_{ij} = \{d_x, d_y, d, r\} \quad (1)$$

式中,  $d_x = x_i - x_j$ ;  $d_y = y_i - y_j$ ;  $d = \sqrt{d_x^2 + d_y^2}$ ;  $r = \frac{2l_{ij}}{l_i + l_j}$ 。其中,  $x, y$  是节点所代表区域的中心点坐标,  $l_{ij}$  是两个区域之间共有的边界长度,  $l_i$  和  $l_j$  分别是区域  $i, j$  的周长,  $r$  是共有边界所占比率;

第 2 类边反映对象之间空间位置关系的先验模糊规则, 且每条模糊规则反映两个模糊子集中, 特定对象之间的二元模糊关系, 比如相邻度规则: 如果对象  $A$  和对象  $B$  的共有边界比率很小, 则输出为 0.8。

一般节点和表示整幅图各节点之间边的定义和上面的定义有所不同, 其第 1 类边由 4 个元素构成, 它们分别是一般节点所表示区域的中心点到图象上

下左右 4 个边界的距离。

$$e_{i0} = \{x_i, x_r, y_i, y_r\} \quad (2)$$

它们的第 2 类边只有方向性模糊规则, 如: 对象  $A$  在图象的右下方。

显然第 1 类边是区域之间的实际空间位置关系, 而第 2 类边是对象之间最有可能的空间关系。

## 2 适应函数的构造

在进行图象解释时, 图象的每一个对象可能的解释都将赋予一个适应函数值, 因为适应函数不仅反映了确定对象在相应模糊子集中的隶属度, 还反映了对象之间实际的空间位置关系和相应的先验空间位置关系的和谐度, 其隶属度越大, 和谐度越好, 适应函数值就越大; 反之亦然。上述隶属度函数可由基于区域特征的模糊分类器给出, 且两个对象之间的和谐度就是模糊规则的输出, 其适应函数是对象空间位置和谐度和对象模糊隶属度的乘积, 而被解释对象的模糊隶属度则是两个相关对象的模糊隶属度的平均。

$$f = \sum_{j,k} y_{jk} \mu_{jk} \quad (3)$$

$$\mu_{j,k} = \frac{\mu_j + \mu_k}{2} \quad (4)$$

其中,  $\mu_j$  和  $\mu_k$  分别是两个相关对象  $j, k$  在各自模糊集中属于特定物体的隶属度;  $\mu_{j,k}$  是平均模糊隶属度;  $y_{j,k}$  是对象空间位置和谐度;  $f$  是适应函数, 它是图中所有相关对象适应度的总和。

## 3 基因搜索

基因搜索是一种仿自然的用于确定全局最优的方法<sup>[6]</sup>, 它首先随机产生一组解释结果, 然后通过复制、交叉和变异操作, 逐渐演化得到正确的解释。针对不同的问题, 复制、交叉和变异 3 种算子的设计差别很大。在下面的讨论中, 一个基因串表示一种解释结果, 基因串中每个基因表示一个节点, 而基因种群则表示一组解释结果。

### 3.1 初始种群的产生

初始种群产生采用随机均匀选择的方法。因为种群的大小决定基因搜索的收敛速率, 即种群越大, 寻找最优解需要的遗传代数相应较少, 但每一代的计算量却很大; 而种群过小, 寻优的概率就降低, 所以要折衷考虑, 可将种群大小设置为与模糊图的节点数相等<sup>[5]</sup>。

### 3.2 复制

基因复制时,可按照各个解释结果(基因串)的适应函数值来决定它们在下一代中出现的概率,这样,适应性好的基因串,复制的可能性就大,反之就小.这里适应函数值  $f_i$  和复制概率  $P_i$  成线性关系

$$P_i = \frac{\alpha f_i}{\sum_{j=1}^M f_j} \quad 0 < \alpha < 1 \quad (5)$$

其中,  $M$  是基因种群的大小,  $\alpha$  是复制在遗传中所占的比率.然后利用轮盘算法产生随机的保留基因串.

### 3.3 交叉

交叉操作可以融合现有的部分最优基因串,以生成一个适应性更好的基因串,但一般交叉操作随机决定两个基因串交换位置的方法不适用于模糊图,因为它不仅造成基因串节点已经是最优的部分被强行拆开,而且将造成密切相关的子图被拆分.

这里提出一种新的交叉方法,即首先按图的特征把图切割为两个子图,而不是随机切割;然后再进行子图的交叉互换,而理想的切割应该使两个子图之间连接边的容量最小,即两个子图的相关性最弱,在图论中就是寻找无向图的最小割切问题<sup>[7]</sup>.假设  $V_1$  和  $V_2$  是  $V$  的两个子集,且  $V = V_1 + V_2$ , 对于一个端点在  $V_1$  上,而另一端点在  $V_2$  上的所有边的集合就是一个割切,而割切的容量就是割切中边的强度之和.由于如今寻找最小割切尚没有一个有效的快速算法,尤其在需要控制两个子集数量大致均衡的环境下,因此本文提出一种比较简单的次最优的图切割方法,其算法如下:

(1) 随机选择两个进行交叉操作的基因串,也就是选择两个具有特定对象的图,首先把它们的边强度相加,以构成一个新的图,然后对新图进行分块.

(2) 选择图中强度最大的边,把它的两个端点作为初始的节点子集  $V_1$ .

(3) 设定一个强度阈值  $\theta$ ,判断一个端点在  $V_1$  上,而另一端点不在  $V_1$  上的全部边,如果边的强度大于  $\theta$ ,那么该边的另一端点就加到  $V_1$  中,形成新的子集  $V_1$ ;如果所有边的强度均小于  $\theta$ ,那么分块结束.  $V_1$  的补集就是  $V_2$ .

(4) 重复第3步,直到分块结束.

这种图切割方法基本上保证了强连接节点没有被拆分,而且保留了一定的随机性.另外,强度阈值一般选择  $0 \leq \theta \leq 0.5$  比较合适.最后按照划分的结果来对两个基因串  $V_i$  和  $V_j$  进行交叉互换,交叉的

概率为  $\beta$ .

### 3.4 变异

变异是增强基因遗传的随机性,它按变异概率  $\gamma$  随机选择基因种群中的部分基因串作为操作对象,并随机改变这些基因串中一定数量的基因.

### 3.5 基因搜索过程

(1) 建立初始的搜索种群.

(2) 对当前的种群进行复制、交叉和变异操作,形成新的种群.

(3) 重复第2步的操作,直到满足收敛条件,本文采用的收敛条件是:①迭代次数超过种群大小的20倍;②在10次连续迭代中,最好的基因串没有发生变化.当两个条件同时满足时,则操作结束.

(4) 选择最好的基因串作为图象解释结果.

在一般情况下,复制、交叉和变异3种操作的比例大致为  $0.4 : 0.5 : 0.1$ ,当然也可随着遗传代数的增加,自适应地修改3种操作的比例.由于本文在设计交叉算子的时候,考虑了强连接节点不被拆分,这就避免了许多盲目和无效的搜索,在实验中,还发现采用这种交叉算子比用随机交叉算子收敛速度提高了5倍以上,加上收敛条件选择的合理性,基因算法的收敛性是可以保证的.

如果存在非单一性对象的区域,本文提出的基因搜索算法也可以在不改变基本算法结构的情况下,有效处理这一问题.这里,非单一性对象是指把由多个物理对象构成的子区域节点解释成其中的任意一个物理对象都不会违反先验的空间模糊关系规则,即都有较好的适应性,这样搜索结果就有多个最优或次最优的解释结果,把这些结果综合起来,就可得到一个更合理的解释,其综合过程如下:

(1) 在最后一代种群中,首先选择一组具有较高适应函数值的基因串,并使这一组中基因串的适应函数值最大和最小值的差别小于设定的阈值,然后选择适应函数值最大的基因串作为初始综合结果.

(2) 按适应函数值的大小次序,选择下一个基因串,把两个基因串合为一个,即部分节点有多个物理对象,如果物理对象没有空间位置关系的明显矛盾,那么就构成一个新的综合结果;否则综合解释结果不变.

(3) 重复第2步,直到一组基因串均被检查一遍,此时得到的综合解释结果就是最后的解释结果.

## 4 实验分析

实验1 这是一个人造数据的仿真实验,假设

有 9 个需要识别的分割区域(图 1), 其中可能的物理类别有 6 个, 在没有空间位置信息的情况下, 各个区域相对于每一个类别的模糊隶属度如表 1 所示, 而模糊空间位置关系规则如表 2 所示.

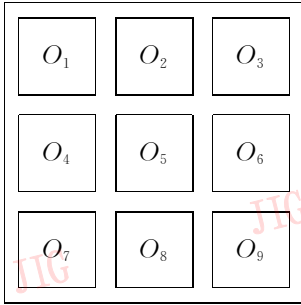


图 1 分割区域

表 1 初始分类模糊隶属度

	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$
$O_1$	0.5	0.5				
$O_2$		0.5	0.5			
$O_3$			0.5	0.5		
$O_4$			0.1	0.8	0.1	
$O_5$				0.5	0.5	
$O_6$			0.1		0.2	0.7
$O_7$	0.6	0.4				
$O_8$				0.2	0.8	
$O_9$					0.4	0.6

表 2 模糊空间位置关系规则

	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$	整图
$A_1$		左侧				距离远	左侧
$A_2$	—		相邻		上方		上方
$A_3$	—	—				上方	右侧
$A_4$	—	—	—			左侧	
$A_5$	—	—	—	—	相邻		
$A_6$	—	—	—	—	—		右侧

—表示与上半部呈对称关系.

通过基因搜索, 可得到如下 4 个具有最大适应函数值的基因串:

- $O_1$   $O_2$   $O_3$   $O_4$   $O_5$   $O_6$   $O_7$   $O_8$   $O_9$
- (1)  $A_1$   $A_2$   $A_3$   $A_4$   $A_5$   $A_6$   $A_1$   $A_5$   $A_6$
  - (2)  $A_1$   $A_2$   $A_3$   $A_4$   $A_4$   $A_6$   $A_1$   $A_5$   $A_6$
  - (3)  $A_1$   $A_2$   $A_3$   $A_4$   $A_5$   $A_6$   $A_1$   $A_5$   $A_5$
  - (4)  $A_1$   $A_2$   $A_3$   $A_4$   $A_4$   $A_6$   $A_1$   $A_5$   $A_5$

最终的解释结果为

$$A_1 A_2 A_3 A_4 A_4/A_5 A_6 A_1 A_5 A_5/A_6$$

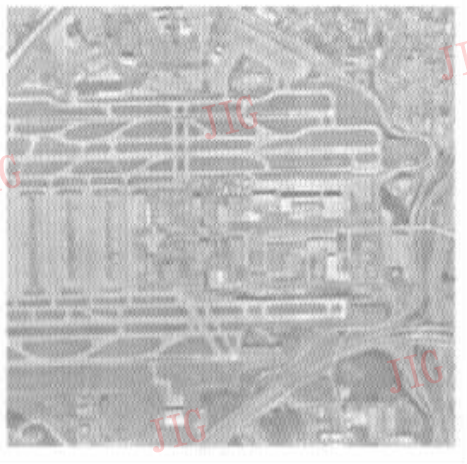
该实验结果表明, 在初始分类非常模糊和不准确的情况下, 通过表示对象之间先验空间位置关系的模糊规则, 即可以对初始结果进行修正和重新确认, 并能得到正确的解释结果.

实验 2 这是一个航空照片自动解释的实验, 该照片是一幅美国亚特兰大机场和周边环境的航空照片. 它由 6 幅  $225 \times 150$  的小图拼成如图 2(a) 所示的  $450 \times 450$  的大图. 其实验步骤为:

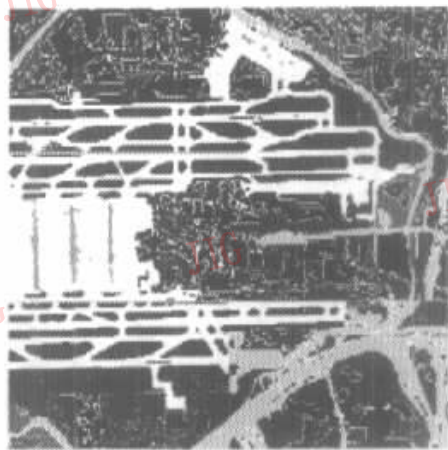
(1) 采用文献[8]中提出的基于图象全局和局域信息的区域分割方法, 把图象分成需要识别的子区域, 同时计算每个子区域的如下统计/几何特征: ①区域大小, ②区域的平均灰度, ③区域的纹理特征, ④区域的长宽比, ⑤区域的紧凑度.

(2) 采用文献[9]中提出的基于 ID3 方法的模糊神经网络分类器来得到各子区域关于识别物体的隶属度, 其中, 针对机场目标需要解释的物体有机场跑道/停机坪、停靠的飞机、重要的公路、草地、树林和建筑物; 然后提出解释物体空间位置关系的模糊规则, 如“飞机都在跑道或停机坪上”, “跑道中间夹杂大块草坪”, “机场附近必有重要的公路”等.

(3) 利用基因搜索来完成整个图象的解释, 如图 2(b) 所示. 由图 2(b) 可以看出, 解释结果是好的. 这里需要说明的是, 由于分割后形成的子区域数量很大, 因此为了节约时间, 基因搜索过程中, 只对大的区域进行操作, 而小区域则采用合并到邻近大区域的方法来完成解释. 如果单纯采用文献[9]中的模糊分类器来得到各子区域关于识别物体的模糊隶属度, 那么就无法得到确定的解释结果; 如果采用简单的取最大隶属度的方法来得到确定解释, 则发现其实验结果是比较差的, 这说明基于先验规则和基因搜索推理的方法是一种有效的改进图象解释性能的方法.



(a) 机场及周边地区航空照片



(b) 自动解释图

图 2 美国亚特兰大机场和周边环境航片及其自动解释结果

■ 建筑物	■ 树	■ 道路
■ 草地	■ 飞机	□ 机场跑道

### 5 结束语

大家知道,统计/几何特征和对象关系特征这两类主要的图象解释信息是互为补充的,而本文则通过基因搜索算法中的适应函数综合了这两类解释信息,其中基于统计/几何特征的解释信息由模糊分类器处理,而关系特征则由模糊图推理处理。由于本文方法充分利用了这两类信息以及基因搜索算法的优点,从而为复杂图象解释系统的构造提供了新的途径。

图象解释是一个非常困难的人工智能问题,目前的理论和实践研究认为,如果要在实际应用中取得较好的效果,那么,分割过程和解释过程不应是顺序关系,而应是自上而下和自下而上的交互作用关系。对于基因搜索来说,就是一个变结构的问题,虽然从理论上讲,基因算法已提供了解决变结构问题的能力<sup>[6]</sup>,但关键在于如何设计基因操作算子,尤其是在基因搜索过程中,如何实现对分割区域的重新定位将是下一步的研究课题。

### 参 考 文 献

- 1 Dillon C, Caelli T. Learning image annotation: The CITE system[J]. Journal of Computer Vision Research, 1998,1(2): 90~121.
- 2 Matsuyama T, Hwang V. SIGMA: A knowledge-based aerial image understanding system[M]. New York: Plenum Press, 1990.
- 3 王润生. 图象理解[M]. 长沙: 国防科技大学出版社, 1995.

- 4 Kumar V P, Desai U B. Image interpretation using bayesian networks [J]. IEEE Trans. Patt. Anal. and Mach. Intel., 1996,18(1):74~77.
- 5 Cross A D J. Inexact graph matching using genetic searching [J]. Pattern Recognition, 1997,30(6):953~970.
- 6 Goldberg D E. Genetic algorithms in search, optimization and machine learning[M]. Reading, MA: Addison-Wesley, USA, 1989.
- 7 卢开澄, 卢华明. 图论及其应用[M]. 北京:清华大学出版社, 1995.
- 8 Qian Y T, Zhao R C. Image segmentation based on combination of global and local information [A]. In: ICIP'97 [C], Santa Barbara, CA, USA, Oct, 26-29, 1997:204~207.
- 9 Qian Y T, Xie W X. Constructing fuzzy neural classifiers by ID3 algorithm [A]. In: ICSP'96 [C], Beijing, Oct. 14-18, 1996: 1292~1295.

钱 涛 1968 年生,副教授,1996 年获西安电子科技大学信号与信息处理专业博士学位,1996 年至 1998 年在西北工业大学做博士后,1999 年至 2000 年在加拿大 Concordia 大学做访问科学家。现在浙江大学计算机系从事模式识别、图象处理和计算智能方面的研究。已发表论文近 30 篇。

王 琦 1970 年生,助理工程师,1997 年获计算机专业工学学士学位,1992 年至 1998 年在陕西省测绘局数据处理中心从事测绘数据和图象分析工作。现在浙江大学计算机系机房工作,从事计算机实验教学和图象数据分析研究。