

图象的符号表述方法

王 绍 霖

摘 要 图象的符号表述方法是用计算机对视觉图象进行描述的基本环节,在本文中,讨论各种不同的符号表述方法,它们是象素特征、形状特征、几何结构、句法关系等,这些特征可以分成属性特征和非属性特征两大类。与之相应,测度物体或子图象的相似程度的距离也有属性距离和非属性距离两类。还介绍了语义网和松弛标记法,这两种方法与人类视觉处理过程更相似些。

一、引 言

数字图象处理经过近廿多年的研究和发展已广泛地进入工业、医学等实用领域,已进入了感觉和认识的阶段,即计算机视觉的阶段。计算机视觉就是用计算机模拟人眼视觉系统对外界世界进行理解和识别,因此它与人工智能有着非常密切的关系,也可以说是属于人工智能中非常重要的一个部分。在数字图象处理这一新的阶段,如何用符号来表述一个图象是使计算机与视觉世界进行联系的必要手段。只有用简洁而有效的符号(包括数字)来表述图象之后,才有可能用计算机对图象进行理解和认识。但是视觉图象是非常复杂的,因此图象的符号表示也是非常困难的,到目前为止并没有得到完善的解决。

符号表述的基本原则:一是简洁,这样计算机才能以方便的算法来处理图象;二是有效,也就要能足够有效地表达图象中复杂的信息。在图象信息进入三维景象领域时,由于物体的互相遮掩作用,在分析这些图象时必须用推断的方法,即利用对物体的过去的知识和联想推断对图象中的物体进行比较完整的描述。本文对图象的符号表述的当前的一些方法进行介绍和讨论。

二、图象的特征表述方法

图象的特征表述方法是最常用的,也是最基本的一种图象表述方法,在以往的图象处理和分析中都采用这种方法,它的基本概念是用一组数字或符号来表示一个图象的象素或一个图象中的区域(例如目标、物体及其集合),通常用特征矢量 X 表示:

$$X = \{x_1, x_2, \dots, x_n\}$$

这种特征矢量的表示方法对于图象模式分类识别是很适用的。特征的内容则由于图象分析领域的开拓而不断具有新的含意,由简单到复杂。下面分几类介绍这些特征量的含意。

1. 象素特征

以每一个象素作为表述单元,象素的特征比较简单,不外乎灰度值(f_g)、颜色(f_c)或波段(包括红外、微波、热辐射等)($f_{g1}, f_{g2}, \dots, f_{gn}$),以及统计参数,如均值 m 、方差 σ 等。当然这时把每一象素当成一个独立的随机变量对待,但是这些统计参数并不容易获得。而且图象各象素的统计值是空间移变的,要测量获得这些参数是很困难的。因此有时只能以直方图来代替统计参数,这当然是很不相称的。象素特征在卫星遥感图象的分类和识别是经常使用的,这是由于在遥感图象中,一个象素就代表约 79×79 米的面积的地面。但为了分类更精确有时必需利用象素间的关系信息。

2. 形状特征

在大多数情况下,被描述和辨识的对象并不是一个象素点而是一个二维的区域。如在文字识别、细胞分类、景物分析、侦察照片判读、军事目标的识别跟踪等,在这种情况下单个的象素特征已不再是有效的。一个二维区域的象素集合,以及对这个象素集合的描述将更为复杂,二维区域的形状也是这种描述的恰当的方法之一。

要提取一个目标区域的形状特征参数,首先要对图象进行分割处理,即把图象中的目标物体和背景明确地区分开,把目标物体所占有的象素以“1”标示,把属于背景的象素以“0”标示。这种分割也称为图象的二值化,分割后的图象象素数目不变,但只有0和1两种标志。称为0—1图,或标志图,分割的方法常用的有:

- 以灰度为依据的门限分割法
- 以直方图为依据的门限分割法
- 以物体边缘线为依据的分割法
- 边缘线的统计检测法
- 区域溶合和分裂技术(Merge and split)
- Hough变换提取边缘线
- 以纹理(texture)为依据的分割法

由于分割质量的好坏直接影响到所提取物体的形状,因此分割技术是测量物体形状特性的关键,必须按图象的性质作出恰当的选择,这将在另文中讨论。

常用的形状特征参数有:

a) 周长 L

对数字化图象中目标物体的周长计算必须对已分割的0—1图用积分几何的方法。例如当目标区域为单连通区域时

$$L = \pi \cdot J_{\alpha\alpha} \quad (1)$$

$$J_{\alpha\alpha} = \frac{1}{2\pi} \int_0^{2\pi} J(\alpha) d\alpha$$

式中 $J(\alpha)$ 为该区域在 α 方向上的投影宽度。

b) 圆度 K

当目标区域由平面简单闭曲线组成时

$$K = 4\pi s / L^2 \quad 0 \leq K \leq 1 \quad (2)$$

式中 s 为目标区域的面积, 这可以用简单的对分割后目标区域中的象素进行计数得到。圆度是一个运动不变量, 即它是位移、旋转、大小不变的。当 $K \rightarrow 1$ 时, 物体近于圆;

$K \rightarrow 0$ 时物体为一细线; $K = \frac{\pi}{4}$ 为正方形; $K = \frac{\sqrt{3}}{9} \pi$ 时为三角形。

c) Cauchy-Crofton 方向数

$$C(\theta) = \frac{D(\theta)}{J(\theta)} \quad (3)$$

式中 $J(\theta)$ 表 θ 方向的投影宽度; $D(\theta)$ 为广义方向直径, 又称方向投影总和。当区域为一凸闭曲线时, 广义方向直径 $D(\theta)$ 即方向投影宽度, (3) 式等于 1。当物体在 θ 方向上有凹时, $C(\theta) > 1$, 凹得越深 $C(\theta)$ 越大; 凹处越多, $C(\theta)$ 也越大。故 C 又称为方向凹度。

d) 其它形状参数

D_{\max} —— 主轴长度; θ_{\max} —— 主轴方向;

$Z = \pi x \frac{1}{2\pi} \int_0^{2\pi} J(\alpha) d\alpha$ —— 最小凸包周长;

$C = L/Z$ —— 整体参数;

$x = M - H$ —— 欧拉数; M 为连通物体数, H 为物体的孔数;

x_0, y_0 —— 物体的几何中心;

I_0 —— 中心惯性矩; I_1 —— 主轴惯性矩。

e) 物体边缘线描述方法 —— 链码

Freeman 链码为表述物体边缘线的方法。

3. 统计特征

一个已经被分割出来的目标物体的区域中的灰度往往不是同一的, 但它具有一定的散布特性, 因此二维统计特征往往是一种有用的目标特征表述工具。二维统计特性是很难测量的, 因此一般用二维直方图来表示。二维直方图也称空间灰度关系法, 一般用一个空间灰度关系矩阵来表示

$$S_\theta(d) = S_\theta(i, j/d) \quad (4)$$

上式表示从一灰度为 i 的象素到灰度为 j 的象素之间距离为 d 的象素对的数目, 如图象灰度分层数目为 M , 则 $S_\theta(d)$ 为 $M \times M$ 对称矩阵。 θ 为 d 的方向。某一个方向 θ 的距离 d 就有一个矩阵, 这是不便于表述的。如取 θ 为各个方向的平均值, 则得一个矩阵

$$S(d) = \sum_{\theta} S_\theta(i, j/d) = S(m, d) \quad (5)$$

$$m = |i - j|$$

如取 θ 为 0 ; $\frac{\pi}{2}$; π ; $\frac{3\pi}{2}$ 四个值, 则有四个矩阵。

这称为差值矩阵。当然这种矩阵还可以有许多各种不同的二维灰度关系表示方法, 也可以用关系与灰度共生矩阵, 差值与灰度共生矩阵等各种不同的二维统计矩阵表示法。特征值可以用对矩阵的一些测量值来表示。例如

$$\text{能量} \quad E[S_{\theta}(d)] = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} [S_{\theta}(i, j/d)]^2$$

$$\text{熵} \quad H[S_{\theta}(d)] = - \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} S_{\theta}(i, j/d) \log S_{\theta}(i, j/d)$$

$$\text{相关量} \quad C[S_{\theta}(d)] = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} (i - \mu_x)(j - \mu_y) S_{\theta}(i, j/d) / \sigma_x \sigma_y$$

$$\text{局部一致量} \quad L[S_{\theta}(d)] = \sum_i \sum_j \left(\frac{1}{1 + (i-j)^2} \right) S_{\theta}(i, j/d)$$

$$\text{惯性矩} \quad I[S_{\theta}(d)] = \sum_i \sum_j (i-j)^2 S_{\theta}(i, j/d)$$

$$\text{绝对值} \quad V[S_{\theta}(d)] = \sum_i \sum_j |i-j| S_{\theta}(i, j/d)$$

$$\text{自相关} \quad A[S_{\theta}(d)] = \sum_i \sum_j ij S_{\theta}(i, j/d)$$

式中

$$\mu_x = \sum_i i \sum_j S_{\theta}(i, j/d)$$

$$\mu_y = \sum_j j \sum_i S_{\theta}(i, j/d)$$

$$\sigma_x^2 = \sum_i (i - \mu_x)^2 \sum_j S_{\theta}(i, j/d)$$

$$\sigma_y^2 = \sum_j (j - \mu_y)^2 \sum_i S_{\theta}(i, j/d)$$

其它各种矩阵也可作类似的特征值计数, 即把一个矩阵经过计算变成几个参数。

三、复杂图象的分级表述方法

需要描述的图象非常复杂, 这是指图象中目标物体的类别 m 和每个目标类的特征数 n 都非常大。简单地用特征矢量来表述目标物体使计算量很大, 而且效率很低。这种情况下可用分层表述方法, 即先把图象 p 划分成 k 个子图象, 每一个子图象 p_j 还可划分成更小的子图象, 我们把最后一级子图象称为图象基元, 而每一个基元则可以用一特征矢量来表述。这样由于基元类别数很小, 使计算量相对减少, 而各基元和子图象之间可用句法结构来表示它们的关系 (图 1)。

设两个子图象或基元 X 和 Y , 它们之间有某种关系 R 联系着, 形成一个图象或子图象, 可写成

$$S \rightarrow XRY$$

一般地讲 X 、 Y 和 R 都可以用它们各自的特征矢量来表示, 在这里这种特征也称属性。

$$A(X) = \mathbf{X} = [x_1, x_2, \dots, x_n]^t$$

$$A(Y) = \mathbf{Y} = [y_1, y_2, \dots, y_m]^t$$

$$A(R) = [r_1, r_2, \dots, r_l]^t$$

$$A(S) = \Phi[A(X), A(Y), A(R)]$$

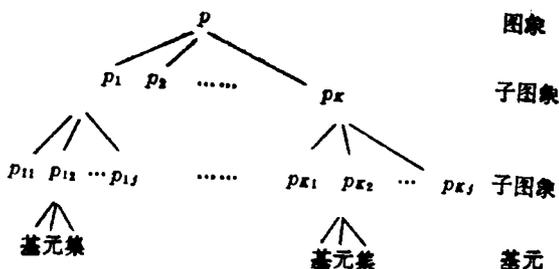


图 1

A 表示属性， Φ 为函数。下面举例说明以上的概念。

例1：设 X 是一属性为长度 l_x 、方向角为 θ_x 的直线段， Y 是长度为 l_y 、方向角为 θ_y 的直线段。把 X 和 Y 连接起来以形成一个图形 s ，则 s 的长度和方向可按下式计算：

$$l_s = l_x + l_y$$

$$\theta_s = \tan^{-1}[(l_x \sin \theta_x + l_y \sin \theta_y) / (l_x \cos \theta_x + l_y \cos \theta_y)]$$

因此属性表示为

$$A(X) = (l_x, \theta_x); A(Y) = (l_y, \theta_y); A(s) = (l_s, \theta_s)$$

例2：属性形状文法

如有一条边缘线它可以分解成一些曲线段 c ，每一曲线段的属可表示为

$$A(C) = (\vec{C}, L, \theta, Z)$$

\vec{C} 为曲线段的矢量长度（连接曲线头尾的矢量长度）， L 为曲线总长， θ 为曲线的总改变角， Z 为的对称性度量。

如有两条曲线它们之间的夹角为“ a ”，把 C_1 和 C_2 连接起来形成线条 N 。

$$N \rightarrow C_1 \text{ cat } C_2$$

$$A(N) = A(C_1) +_a A(C_2)$$

上式的意义是 $\vec{C}_N = \vec{C}_1 + \vec{C}_2$ ； $\theta_N = \theta_1 + a + \theta_2$ 。如 S 线条由 N_1 和 N_2 组成，则

$$S \rightarrow N_1 \text{ cat } N_2$$

$$A(s) = A(N_1) +_a A(N_2)$$

这里可以明显地看到分层结构。

例3：PDL 图象描述语言

PDL 描述语言也是用来描述复杂的曲线或细长形物体的，这里基元是一个简单的长形物，它可用两个点来标记，一个称为头，一个称为尾。一个基元只能以其头或尾连接到其它的基元上。这种方法对文字、图形等进行表述特别适用。有两种连接算子和一个反转算子 \sim 。

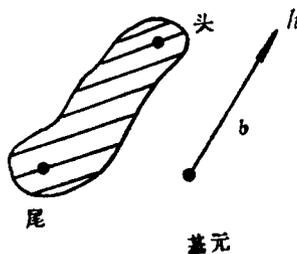
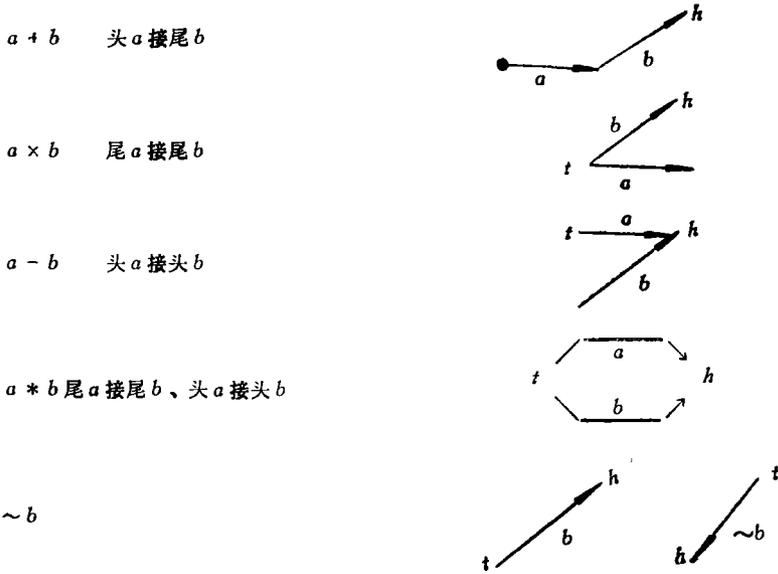


图 2



这四个算子可用 $CAT(+, \phi)$; $CAT(-, \phi)$; $CAT(\times, \phi)$; $CAT(*, \phi)$ 来表示加入到尾性特征中去。产生PDL描述句子的文法是上下文无关的属性文法

$$G = (V_N, V_T, p, s)$$

其中

$V_N = \{s, sL, \phi_b\}$ 为非终止符;

$V_T = \{b\} \cup \{+, \times, -, *, \sim, /, (,)\} \cup \{l\}$ 为终止符。

算子 $/$ 表示头和尾处在任意位置, b 为任意基元, l 是指示标号的, 它用来交叉地参考描述中的表示式 S 。 $S \in V_N$ 为起始符, p 是再写规则, PDL规则是

$$\begin{aligned}
 S &\rightarrow b, S \rightarrow (S\phi_b S), S \rightarrow (\sim S), S \rightarrow SL, \\
 S &\rightarrow (/SL), SL \rightarrow S^l, SL \rightarrow (SL\phi_b SL), \\
 SL &\rightarrow (\sim SL), SL \rightarrow (/SL), \\
 \phi_b &\rightarrow +, \phi_b \rightarrow \times, \phi_b \rightarrow -, \phi_b \rightarrow *
 \end{aligned}$$

例如: 用基元 \xrightarrow{a} , \nearrow^b , \searrow^c , \uparrow^d 来表示字母 A , 表示式为 $(b + ((b + c) * a) + c)$

PDL 规则中有两类, 一类是句法法则, 一类是语义法则, 句法法则的形式如上列所示, 是上下文无关的, 可概括为

$$\begin{aligned}
 N &\rightarrow a, N \in V_N \\
 a &= C_1, C_2, \dots, C_k; C_j \in (V_N \cup V_T)
 \end{aligned}$$

语义法则是一个映射 f

$$f : DC_1 \times DC_2 \times \dots \times DC_k \rightarrow D_N$$

也可写成

$$A(N) = f(A(C_1), A(C_2), \dots, A(C_k))$$

这个映射是一个闭合型函数, 它表示 N 的属性可表示成 C_1, C_2, \dots, C_k 的属的函数。当这两类规则同时存在时组成属性文法。当把所有的语义规则去除后就变成上下文无关文

法。这种语义和句法法则的共存可以达到用比较简单的句法结构辅之以一定的语义信息来对复杂的图象模式进行描述。可从下例看到。

例4：一种大家熟悉的三角形集的描述语言 $L = \{a^n b^n c^n / n = 1, 2, \dots\}$ 是上下文敏感语言 G_1 ，其再写规则为：

$$S \rightarrow aSBA; \quad bB \rightarrow bb; \quad S \rightarrow aBA; \quad bA \rightarrow bc;$$

$$AB \rightarrow BA; \quad cA \rightarrow cc; \quad aB \rightarrow ab.$$

如加入属性 l_a, l_b, l_c 和 θ ，它们表示基元 a, b, c 的长度，以及基元连接的角度，则可以变成上下文无关文法 G_2 ，下面用 PDL 记法来表示；以 $CAT(a, b)$ 表示 a 与 b 之间的连接关系， $CAT(a, b) = (+, \theta)$ ，再写规则为

$$S \rightarrow ABC; \quad CAT(S, A) = CAT(C, a) = (+, 120^\circ);$$

$$l_S = l_A + l_B + l_C;$$

$$A \rightarrow aA; \quad CAT(A, B) = CAT(a, b) = (+, 120^\circ);$$

$$l_A = l_a + l_A;$$

$$B \rightarrow bB; \quad CAT(B, C) = CAT(b, c) = (+, 120^\circ);$$

$$l_B = l_b + l_B;$$

$$C \rightarrow cC; \quad CAT(C, a) = CAT(C, a) = (+, 120^\circ);$$

$$l_C = l_c + l_C$$

$$A \rightarrow a; \quad CAT(a, b) = (+, 120^\circ); \quad l_A = l_a = 1;$$

$$B \rightarrow b; \quad CAT(b, c) = (+, 120^\circ); \quad l_B = l_b = 1;$$

$$C \rightarrow c; \quad l_C = l_c = 1.$$

上面的表示式中每一行前面部分的规则是句法规则，后面加入的是语义规则。

也可以表示为有限状态文法 G_3 ：

$$S \rightarrow aA; \quad CAT(a, b) = (+, 120^\circ); \quad l_S = l_a + l_A;$$

$$A \rightarrow bB; \quad CAT(b, c) = (+, 120^\circ); \quad l_A = l_b + l_B;$$

$$B \rightarrow C; \quad CAT(c, a) = (+, 120^\circ); \quad l_B = l_C;$$

$$l_a = l_b = l_c = n, \quad n = 1, 2, \dots$$

从上面的例子可以看到用加入较为复杂的语义可使一个模式的句法表述从上下文敏感文法变为上下文无关或有限状态文法。

四、距离测量法

由于图象中存在噪声，以及图象中被摄取的物体本身的变化，或环境因素的变化，使得在图象中对目标物体区域的特征测量以及关系结构都不能在各次不同的时间、地点、场合达到完全相同的程度，因此必须进行相似性度量。距离度量是一种简便而有效的相似性度量方法，被经常用在各种不同的模式分析和分类问题中。从上一节可以看到对图象进行描述的特征有属性特征和非属性特征两种。在进行距离计算时也有两种距离算法。属性距离又可分成非二进和二进两种情况。

1. 属性距离

两个属性特征矢量串 X_i 和 X_j 之间的距离 d 是一个 $X \times X$ 域的实对称函数, 它应服从距离三公理, 即

- i) $d(i, i) = 0$ 自反性
- ii) $d(i, j) = d(j, i)$ 对称性
- iii) $d(i, j) \leq d(i, k) + d(k, j)$ 三角不等式

对于非数量参数的距离, 第iii)条改

$$\delta(i, j) \leq \sup[\delta(i, k), \delta(k, j)]$$

称为超参量距离。

常用的距离测度方法有以下几种:

- i) 明可夫斯基距离 $d_1(X_i, X_q) = \left[\sum_{j=1}^n |X_{ij} - X_{qj}|^{1/\lambda} \right]^\lambda$
- ii) 弯曲距离 $d_2(X_i, X_q) = \sum_{q=1}^n |X_{ij} - X_{qj}| / |X_{ij} + X_{qj}|$
- iii) 切比雪夫距离 $d_3(X_i, X_q) = \max_j |X_{ij} - X_{qj}|$
- iv) 欧氏距离 $d_d(X_i, X_q) = (X_i - X_q)' Q (X_i - X_q)$

Q 为 $-n \times n$ 正定距离。

- v) Mahalanobis距离 $d_s(X_i, X_q) = (\det W) (X_i - X_q)' W^{-1} (X_i - X_q)$
- vi) 相关 $d_6(X_i, X_q) = \left[\sum_{j=1}^n (X_{ij} - \bar{X}_j)(X_{qj} - \bar{X}_j) \right] / \left[\sum_{j=1}^n (X_{ij} - \bar{X}_j)^2 \sum_{j=1}^n (X_{qj} - \bar{X}_j)^2 \right]^{1/2}$
- vii) $d_7(X_i, X_q) = \sum_{j=1}^n W_j |X_{ij} - X_{qj}|$
- viii) $d_8(X_i, X_q) = \sum_{j=1}^n \frac{1}{X_{0j}} \left[\frac{X_{ij}}{X_{i0}} - \frac{X_{qj}}{X_{q0}} \right]^2$

$$\text{这里 } X_{0j} = \sum_{i=1}^m X_{ij}, \quad X_{i0} = \sum_{j=1}^n X_{ij}$$

2. 二进参数距离

当 X_i 和 X_q 都是二进制的参数串时, 令 a, b, c, e 分别是:

- a 为 $X_{ij} = 1$ 和 $X_{qj} = 1$ 同时出现的次数;
- b 为 $X_{ij} = 0$ 和 $X_{qj} = 1$ 同时出现的次数;
- c 为 $X_{ij} = 1$ 和 $X_{qj} = 0$ 同时出现的次数;
- e 为 $X_{ij} = 0$ 和 $X_{qj} = 0$ 同时出现的次数。

则距离测度也可以有许多形式, 常用的有

- $d_1(X_i, X_q) = a / (a + b + c + e)$
- $d_2(X_i, X_q) = a / (a + b + c)$
- $d_3(X_i, X_q) = a / (2a + b + c)$
- $d_4(X_i, X_q) = a / [a + 2(b + c)]$

$$d_s(X_i, X_q) = a/(ae+bc)/[(a+b)(c+e)(a+c)(b+e)]^{\frac{1}{2}}$$

3. 非属性距离

对非属性的关系表述也可以写成一个序列串, 对这些序列串之间的相似程度也可用距离这个概念来表示, 这种距离一般用Levenshtein距离来表示, 它的定义是从 X_i 串改变成 X_q 串所必需的最少的替代、消去和插入的次数。例如有两个句法表述串 X_i 和 X_q ,

$$X_i = cbbabdbb; X_q = cbabdbb$$

其中每一基元 a, b, c, d 都可用一属性矢量 $A(a), A(b), A(c), A(d)$ 表述。这里的Levenshtein距离 $d_L(X_i, X_q) = 3$, 因为

$$X_q = cbabdbb \xrightarrow{\text{替代}} cbabbbb \xrightarrow{\text{替代}} cbabdbb \xrightarrow{\text{插入}} cbbabdbb = X_i$$

如果把所有不同的基元符号都去掉得 $X'_i = cbabb = X'_q$, 但这两个串的属性不一定相等, 但 $A(X'_i) \neq A(X'_q)$ 。这是由于每一个基元测得的属性特征矢量当然可能有差异。因此具有属性和非属性特性的句法表述串之间的距离可表示为

$$d(X_i, X_q) = \alpha d_L(X_i, X_q) + \beta d_A(X_i, X_q)$$

式中 d_A 为属性距离, 可以用去除不同基元符号后的基元串之间的属性距离之和表示:

$$d_A(X_i, X_q) = \sum_{j=1}^k w_j d(A(X'_{ij}), A(X'_{qj}))$$

关于距离的测度方法还有许多, 对不同的对象应采用不同的测度, 以达到简洁而有效的目地。

五、高级关系表述

人的视觉能力是很复杂的, 单靠前面所述的特征和关系表述还远不能把复杂的视觉图象信息用符号表述得与人的视觉能力相同。因而必须进入更高级的视觉过程的研究。高级的视觉过程涉及许多生理学和心理学研究, 有许多联想、推理、判断的过程, 现在这方面的研究才刚开始, 这里不能深入介绍, 本文只在用符号表述这些高级关系上作些介绍。这些高级的视觉过程有一个共同的特点, 那就是这些过程中的关系都比较松散, 在数学上把这些松散的、模糊的过程用关系模型表示出来可以有許多方法, 这里只介绍两种简单的松散关系。

1. 语义结构关系

语义网络是按人类记忆模型提出来的。语义网络可以作为一种数据结构, 一种主题表述, 一种体现各个现实个体之间的重要的关系, 一种推理规律的表述。这种关系可以是比较复杂的语义关系。语义网络可以用图论中的节点和弧线结构来表示图象中的各个目标与它们之间的关系。例如有节点 X 和 Y , 则可用记号 $L(X, Y)$ 表示 X 与 Y 之间的关系, 这种关系可以是简单的存在与否, 也可以函数关系如

$$L(X, Y) = V$$

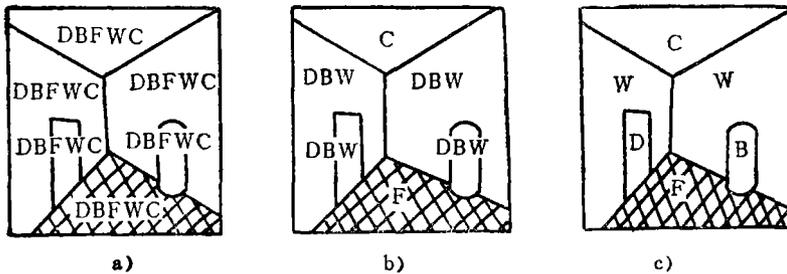
也可以是不能用函数表示的语义关系。语义网络和前节所述关系比较有两点区别: 一是以网络图形形式表示更为复杂的相互关系, 一是这种关系可以用语义形式表示, 比之函

数关系，逻辑关系更为复杂。

2. 松弛标记表述

松弛标记法自1976年 Rosenfeld 提出后已广泛地应用于许多领域，如图象像素标记，图象分割二值化标记，图象物体标记等。在松弛标记法中对象素，物体的各元和整体在开始先不给以一个肯定的标记或类别，而是给以多种标记，每种标记都给以一定的可能性，也称为初始概率。然后，利用各种标记对邻近目标的标记之间的相容关系，用逐次迭代的方法逐步明确这个目标的标记。所有的松弛标记法都必需有四个要素。它们是目标集、目标间的关系集、标记集和约束集。对一个目标的每一个标记可以用“概率”来权衡其确切性，从这一点来讲很类似于用模糊集对事物的表述。约束是对标记所作的限制，它可以是单项进行的，也可以是多项联合的。

现举下图所示的室内景物的例子。室内景物分为六个目标区。共有五个标记，它们分别 D (门)； B (垃圾桶)； F (地板)； W (墙)； C (天花板)。约束条件可分为单项的和关联的，在这里单项的有：



1. 天花板是图内单一的最高区域；

2. 地板必需是方格的；

关联的有：

1. 墙邻接于地板和天花板；

2. 门邻接于地板和墙；

3. 垃圾箱邻接于地板，比门小。

这里只用了非常少的特征就把一个室内景物中的目标区表述清楚了。

松弛标记有离散和连续之分。以上是离散的，即对某一目标的某一标记只有是与非两种状态。连续的标记可用权因子 $p_i(\lambda)$ 表示，目标集为 $a_i, i=1, \dots, n$ ；标记集为 $\lambda_k, \lambda = \lambda_1, \dots, \lambda_m$

$$0 \leq p_i(\lambda) \leq 1; \sum_{\lambda} p_i(\lambda) = 1$$

再定义一个相容系数 $p_{ij}(\lambda/\lambda')$ ，表示对目标 a_i 标记 λ 而另一目标 a_j 标以 λ' 的相容度或条件概率。有

$$\sum_{\lambda} p_{ij}(\lambda/\lambda') = 1; p_{ij}(\lambda/\lambda') = 1, \text{ 若 } \lambda = \lambda', \text{ 否则为 } 0.$$

松弛迭代法可按上面迭代公式进行

$$p_i(\lambda) = \sum_j C_{ij} \left\{ \sum_{\lambda'} p_{ij}(\lambda/\lambda') p_i(\lambda') \right\}$$

或 C_{ij} 为常数因子, $\sum_i C_{ij} = 1$ 。这称为线性迭代法, 但它收敛于一个与初始概率无关的值上, 结果只与相容度有关, 因此与松弛迭代法的本意不尽一致。现在通用的是非线性算法, 其中之一为, 令相容系数为

$$-1 \leq r_{ij}(\lambda, \lambda') \leq 1$$

$$q_i^{(k)}(\lambda) = \sum_j C_{ij} \left[\sum_{\lambda'} r_{ij}(\lambda, \lambda') p_j^{(k)}(\lambda') \right]$$

$$p_i^{(k+1)}(\lambda) = \frac{p_i^{(k)}(\lambda) [1 + q_i^{(k)}(\lambda)]}{\sum_{\lambda} p_i^{(k)}(\lambda) [1 + q_i^{(k)}(\lambda)]}$$

这种算法的收敛性并没有得到充分的证明, 因此可以有许多变形算法。松弛迭代法比较接近于人的视觉功能, 因此在各个方面得到应用。

五、小 结

如何用符号表示一幅复杂的图象是项困难的工作, 在计算机视觉中把图象的表述分为四个等级, 即: 一般图象——数字化以后为 $N \times N$ 个像素点; 分割图象——把背景和目标物明确分开的图象, 仍保持 $N \times N$ 点; 几何结构图象——表示为各个目标物、图象区域或基元的特征及其相互之间简单的几何关系, 已不存在 $N \times N$ 点; 关系图象——表示各目标物之间的高级关系结构。在每一层图象中都要用一定的符号来表述其中的各个元素, 这是用计算机代替人眼视觉功能的最基本的一点。但是由于当前对人的视觉过程的高级阶段, 即人脑的联想记忆、推测、理性推理等过程还理解得很少, 因此这种过程的高级关系描述还很肤浅, 这也正是人工智能这门学科研究的内容。随着第五代计算机研究的进展, 必将得到越来越多的重视。本文只是这方面内容的综述, 不当之处希望大家指正。

参 考 文 献

- [1] K.S.Fu, A General Approach to Picture Analysis Picture Engineering, Springer 1982
- [2] K.S.Fu, 模式识别及其应用, 科学出版社, 1983。
- [3] K.S.Fu, Digital Pattern Recognition, Springer 1976.
- [4] D.H.Ballard, C.M.Brown, Computer Vision, Prentice-Hall 1982.
- [5] 王绍霖, 温殿魁, 数学模式识别(讲义), 国防科技大学, 1985,

The Symbolic Representations of the Pictures

Wang Shaolin

Abstract

The symbolic representations of the pictures are the basic procedure for the computer description of the visual pictures. In this paper, the various methods of the symbolic representations of the pictures are discussed, such as the pixel features, the shape features, the geometric structures and the syntactic relations etc. these features are divided into two groups; attribute and non-attribute, accordingly there are two types of the distances to measure the resemblance of objects or subpictures also; the attribute distance and non-attribute distance. The semantic net and relaxation labeling methods are discussed too. They resemble the human visual processings more.