第10卷 第1期　　自 动 化 学 报　　Vol. 10, No. 1
1984 年 1 月　　ACTA AUTOMATICA SINICA　　Jan., 1984

模式识别的一种词意句法距离度量

戴 汝 为
(中国科学院自动化所)

摘　　要

本文讨论噪声模式和畸变模式的识别问题. 用带有位置坐标的树状文法描述图象，对区域的边界进行分析，建立与产生式相对应的词意规则. 这里既考虑模式的统计特征，又通过词意、句法指导下的变换来描写畸变模式的结构. 在此基础上提出一种包括词意及句法的距离度量，从而用最小距离准则来进行识别.

一、引　　言

用属性文法把统计模式识别与句法模式识别统一起来的努力已经取得了一定的进展[1-4]. 已经知道，一个文法中如果采用位置坐标，则这个文法就超出了乔姆斯基短语结构文法的范畴[5]. 作者在文[2]中曾用向量表示树状文法中树的结点的位置，增加了树状文法的功效，但未解决当模式有噪声或有畸变时的识别问题. 本文讨论用带有位置坐标的树状文法描述和识别包含小区域的图象. 这个文法从二维的角度分析区域的边界，建立与树状文法的产生式相对应的词意规则，并定义了由标准的区域边界集合到畸变的区域边界集合之间的一个词意句法指导下的变换(SSDT)，这个变换的输入文法是有限状态的，而输出文法是上下文无关的，后者反映结构的变化. 在此基础上给出一种既包括词意又包括句法的距离度量，进而根据最小距离准则进行模式识别，这样就把统计方法和句法方法结合了起来.

二、句 法 结 构

考虑一幅图象的构成，分别用 s, t 表示 x, y 轴的单位向量，$N\times N$ 方格上每一方格中心的位置(或坐标)都可以用 s 和 t 的线性组合 v 表示，如图 1 所示. v 处的象素 $\alpha(v)$ (或 $\alpha(i,j)$) 表示 v 处的一个句法符号，$A(\alpha(v))$ 表示 α 的属性.

$$\alpha(v)=\begin{cases}a(v), & A(a)=g,\\ b(v), & A(b)=0,\end{cases}\qquad g\in\{1,2,\cdots,J\}.$$

其中 $a(v)$ 是 v 处的一象素，它的灰度级为 g; $b(v)$ 是空白符号.

用能给出结点位置的树状文法[2]来描绘 $N\times N$ 的图象

本文于 1983 年 7 月 26 日收到.

$$G_t = (\langle V, r\rangle, P, S_0, I).$$

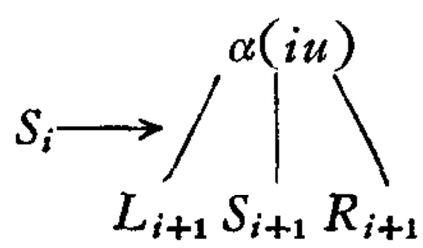
其中

$$V = \{S_i, L_{i+1}, R_{i+1}, i = 0, 1, 2, \cdots, N-1, a, b, A(a) = g, A(b) = 0\};$$

$$r(a) = r(b) = \{0, 1, 3\};\ I = \{s, t, \|s\| = \|t\| = 1, s \cdot t = \pi/2, u = s + t\};$$

s_0 是起始符；

P:



$$L_{i+k} \to \alpha((i+k)t)L_{i+k+1},$$
$$L_{n-1} \to \alpha((n-1)t),$$
$$\alpha \in \{a, b\},\ k = 1, 2, \cdots, n-2-i.$$

以上产生式产生第 i 列中 $\alpha(iu)$ 以下的所有象素.

$$R_{i+k} \to \alpha((i+k)s)R_{i+k+1},$$
$$R_{n-1} \to \alpha((n-1)s),\ \alpha \in \{a, b\},\ k = 1, 2, \cdots, n-2-i.$$

以上产生第 i 行中 $\alpha(iu)$ 右端的象素，另外，

$$S_{n-1} \to \alpha((n-1)u),\ \alpha \in \{a, b\}.$$

文法 G_t 产生一棵与 $N \times N$ 方格上的象素相对的树，如图 2 所示．这样的树表示图象的结构．如果再加上每个象素的属性就能对图象进行完整的描述．另外还可以利用迭加原理用不同的树状文法来描述图象中不同的部分，然后迭加起来．例如分析细胞图象，可以分别用不同的文法来描述胞浆与胞核，再迭加起来就比较有效．

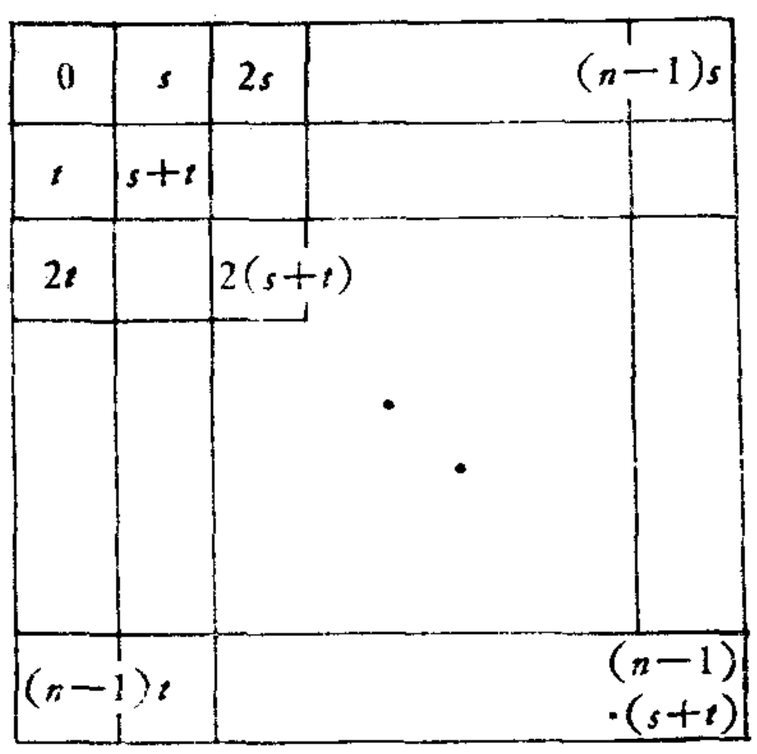


图 1

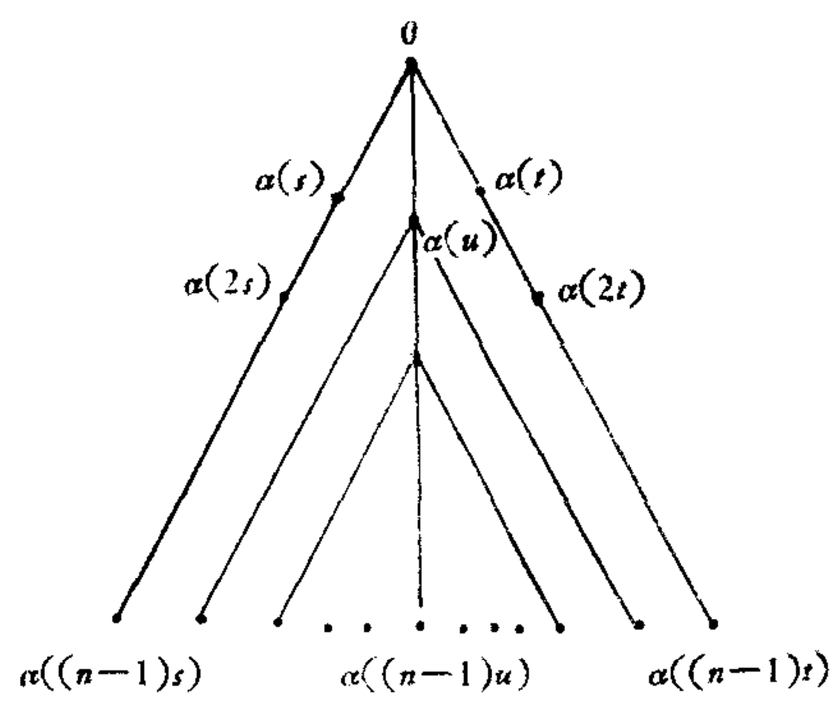


图 2

一个区域的边界所包含的信息对于识别来说是很重要的，为了从图 3 所示的树的结点中决定哪一点是边界点，首先定义点 a 的邻域为 R_a，点 $a(m_1s + m_2t)$ 的邻近的九个点所构成的集为

$$\tilde{R}_a = \{a(m_1s + m_2t),\ a((m_1-1)s + m_2t),\ a((m_1+1)s + m_2t),$$
$$a(m_1s + (m_2-1)t),\ a(m_1s + (m_2+1)t),\ a((m_1-1)s + (m_2-1)t),$$
$$a((m_1-1)s + (m_2+1)t),\ a((m_1+1)s + (m_2-1)t),$$
$$a((m_1+1)s + (m_2+1)t)\}.$$

另外用 R 表示 $N \times N$ 方格上所有 N^2 个点所成的集，把 a 的邻域定义为

$$R_a = \tilde{R}_a \cap R.$$

可以通过下述方法从树的结点中决定边界点：把二维的树的结点标号按照后序（或

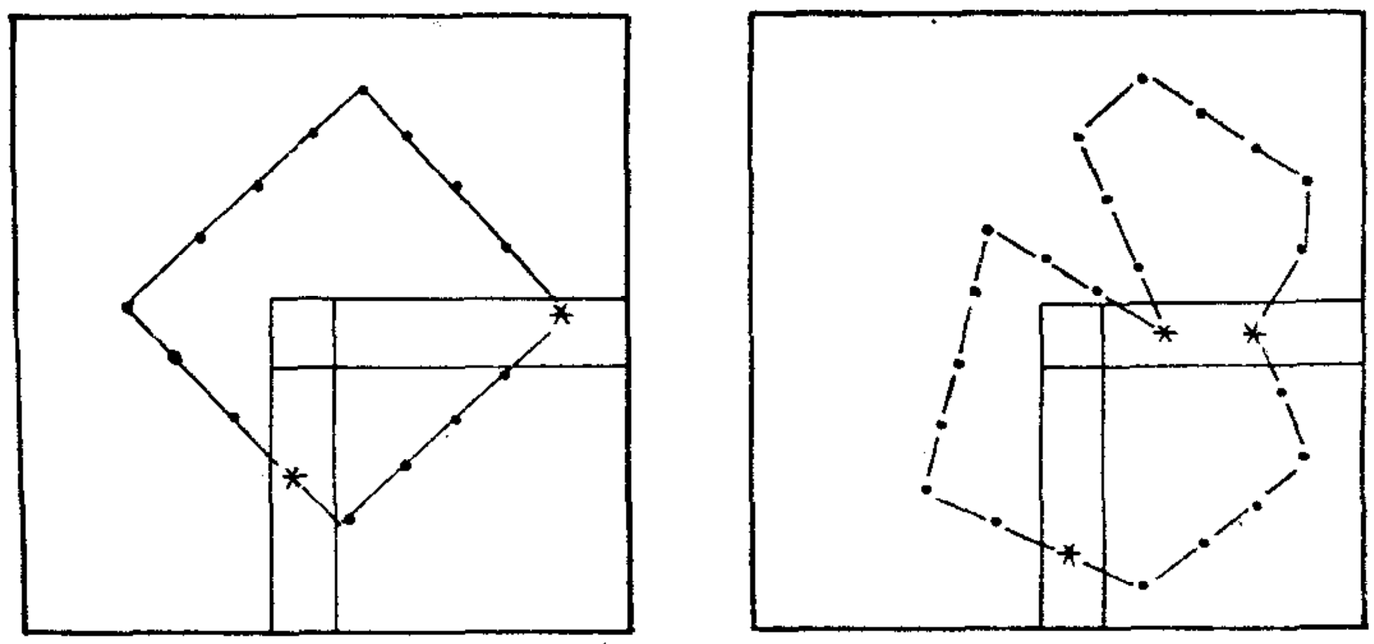


图 3

先序)排列[6],找到一个满足 $A(a)=K$ 的边界点,然后根据连接条件

$$a(v_{i+1})\in R_{a(v_i)},\ A(a(v_i))=K,\ i=1,2,\cdots,n-1$$

排列成一条描述边界点的链 $a(v_1)a(v_2)\cdots a(v_n)$. 令 $\delta_i=v_{i+1}-v_i$ 表示从 v_i 到 v_{i+1} 的向量,上述的链可以表示为 $a(\delta_1)a(\delta_2)\cdots a(\delta_{n-1})$. 这里的连接条件就是一种连接属性.

三、词 意 规 则

树状文法产生树的过程与链文法产生链的过程不同，后者可以通过产生链的导出树(Derivation Tree)来建立词意关系. 根据扩展的树状文法[1]与上下文无关文法的产生式形式上一致的特点，分析包含小区域图象的区域边界特性时，可用上节所述文法 G_t，由 $\alpha(iu)$, L_{i+1}, S_{i+1}, R_{i+1} 的属性得到 S_i 的属性. 如果要用 $A(S_{i+1})$, $A(\alpha(iu))$, $A(L_{i+1})$, $A(R_{i+1})$, $A(S_i)$ 分别表示 S_{i+1}, $\alpha(iu)$, L_{i+1}, R_{i+1}, S_i 的属性,则要确定与产生式对应的词意规则 $A(S_i)=F(A(\alpha(iu)),A(L_{i+1}),A(S_{i+1}),A(R_{i+1}))$. 如果所描述的区域是严格凸的,则 L_{i+1}, R_{i+1} 分别产生一个点;如果不是严格凸的,则可能分别产生若干个连接的或不连接的点,如图 3 所示.为简单起见,这里仅讨论前一种情形.考虑 $S_j\Rightarrow T_j$, $j=0,1,\cdots,n-1$. 按连接条件把边界点排列起来得到一段曲线. 由于词意的计算是倒退的方式,因此用 T_{n-1} 描述边界线段 C_0，用 T_{n-2} 描述边界线段 C_1,……用 T_0 描述 C_{n-1}.

如果用 T_{i+1} 描述 C_i，用 T_i 描述 C_{i+1}，则

$$C_i:\qquad a(v_1)a(v_2)\cdots a(v_i)\ \text{或}\ a(\delta_1)a(\delta_2)\cdots a(\delta_{i-1}).$$

其中 $a(v_{j+1})\in R_a(v_j)$; $\delta_j=v_{j+1}-v_j$, $j=1,2,\cdots,i-1$. C_{i+1} 是在 C_i 的两端分别添加两个点构成的,可表示为

$$C_{i+1}:\qquad a(v_0)a(v_1)\cdots a(v_i)a(v_{i+1})\ \text{或}\ a(\delta_0)a(\delta_1)\cdots a(\delta_{i-1})a(\delta_i).$$

一个区域的边界是一条曲线，一条曲线的属性(或词意)可以用连接曲线尾和头的向量 V、曲线长度 L、曲线的曲率变化 ϕ 以及表示对称性的特征 R 这四个参数,即一个四维向量描述[3]. 这里令

$$A'(C_i)=A'(S_{i+1})=(V_{C_i},L_{C_i},\phi_{C_i},R_{C_i})',$$
$$A'(C_{i+1})=A'(S_i)=(V_{C_{i+1}},L_{C_{i+1}},\phi_{C_{i+1}},R_{C_{i+1}})'.$$

这里的“ ′ ”表示向量的转置向量. 可以求出各分置之间的递推关系[1,3]:

$$V_{C_{i+1}} = \delta_0 + V_{C_i} + \delta_i, \quad L_{C_{i+1}} = l_0 + L_{C_i} + l_i,$$
$$\phi_{C_{i+1}} = \phi_0 + \phi_{C_i} + \phi_i, \quad R_{C_{i+1}} = E_{C_{i+1}} - \phi_{C_{i+1}} L_{C_{i+1}}/2.$$

其中 l_0, l_i 分别表示 δ_0, δ_i 的长度；ϕ_0 为 δ_1 与 δ_0 间的夹角；ϕ_i 为 δ_{i+1} 与 δ_i 间的夹角；$E_{C_{i+1}} = E_{C_i} + \phi_{C_i}(l_0 + l_i)$. 这些递推关系间接地反映了词意规则.

用递推的办法沿树的前沿结点从下往上递推，最终可求得 C_{n-1} 即 S_0 的属性（词意），这个属性反映了整个区域边界的性质. 为了有效地描述区域的性质，还可以用诸如区域的面积、平均灰度等特征作为属性对区域加以描述.

四、词意、句法指导下的变换（SSDT）

当有噪声和干扰时，模式会发生畸变，这可以用误差校正剖析方法来校正[1]. 有噪声或畸变的模式可看成是标准模式加上代换、删除及插入三种误差而成. 对于区域的边界而言，可以认为是标准曲线集合到畸变曲线集合之间的一个以词意、句法为指导的变换.

区域的边界曲线可表示成 $a(v_1)a(v_2)\cdots a(v_n)$. 这里 $A(a(v_j)) = K, j = 1, 2, \cdots, n$，且满足连接条件 $a(v_{i+1}) \in R_{a(v_i)}$，$i = 1, 2, \cdots, n-1$. 其中包括三个基元的一段链 $a(v_{i-1})a(v_i)a(v_{i+1})$. 若有下述三种误差：

1) 删除误差. 把 v_i 处的 $a(v_i)$ 删除，即属性代换

$$A(a(v_i)) = K \Rightarrow A(a(v_i)) = 0,$$

或

$$a(v_{i-1})a(v_i)a(v_{i-1}) \Rightarrow a(v_{i-1})a(v_{i+1}).$$

2) 代换误差. 对 $a(v_i)$ 而言，找一点 $a(v)$ 满足 $A(a(v)) = K$，$a(v) \in R_{a(v_i)}$，$a(v) \in R_{a(v_{i+1})}$，用 $a(v)$ 代换 $a(v_i)$，即

$$a(v_{i-1})a(v_i)a(v_{i+1}) \Rightarrow a(v_{i-1})a(v)a(v_{i+1}).$$

3) 插入误差. 在 $a(v_i)$，$a(v_{i+1})$ 之间插入 m 个点 $a(v^1)$，$a(v^2)$，$\cdots$，$a(v^m)$，满足 $A(a(v^h)) = K, h = 1, 2, \cdots, m$ 及连接条件 $a(v^1) \in R_{a(v_i)}$，$a(v^{j+1}) \in R_{a(v^j)}$，$j = 1, 2, \cdots, m-1$，$a(v^m) \in R_{a(v_{i+1})}$，即

$$a(v_{i-1})a(v_i)a(v_{i+1}) \Rightarrow a(v_{i-1})a(v_i)a(v^1)\cdots a(v^m)a(v_{i+1}).$$

根据上述可能发生的三种误差，定义从标准模式集合到畸变模式集合之间的变换如下：

定义 1. 一个词意、句法指导下的变换方案是一个五元式 $T = (N, \Sigma, \Delta, P, S)$. 其中 N 是一个非终止符的有限集；Σ 是一个有限的输入字母表；Δ 是一个有限的输出字母表；S 是起始符，是 N 中的一个特殊符号；P 是产生式集合. 相邻产生式具有如下形式：

$$N_{i-1} \to (a(v_{i-1})N_i, \quad a(v_i)\bar{N}_i),$$
$$N_i \to (a(v_i)N_{i+1}, \quad a(\xi)N_{i+1}^1 N_{i+1}^i \cdots N_{i+1}^k \bar{N}_{i+1}),$$
$$N_{i+1} \to (a(v_{i+1})N_{i+2}, \quad a(v_{i+1})\bar{N}_{i+2}).$$

令 $\delta_j = v_{j+1} - v_j$，$\tilde{\delta}_j = v - v_j$，则有三种情形：

1) $\xi = v_i$，$N_{i+1}^1 \to \lambda$（零链），$N_{i+1}^2 \to \lambda, \cdots$，$N_{i+1}^k \to \lambda$，

$$a(v_i) \Rightarrow b(v_i)，删除.$$

2) $\xi = v$, $N_{i+1}^1 \to \lambda$, $N_{i+1}^2 \to \lambda, \cdots$, $N_{i+1}^k \to \lambda$. v 满足

$$\mathrm{CAT}(\delta_{i-1}, \delta_i) \Rightarrow \mathrm{CAT}(\tilde{\delta}_{i-1}, -\tilde{\delta}_i)$$ (符号 CAT 见文[6]).

3) 在 $a(v_i)$ 与 $a(v_{i+1})$ 之间插入m个点，$1 \leqslant m \leqslant k$, $\xi = v_i$,

$$N_{i+1}^1 \to a(v^1),\ N_{i+1}^2 \to a(v^2), \cdots,\ N_{i+1}^m \to a(v^m),$$
$$N_{i+1}^{m+1} \to \lambda,\ \cdots,\ N_{i+1}^k \to \lambda.$$

其中

$$a(v^1) \in R_{a(v_i)};\ a(v^j) \in R_{a(v^{j-1})},\ j = 2, \cdots, m-1;\ a(v^m) \in R_{a(v_{i+1})}.$$

令 $\delta^i = v^{i+1} - v^i$，连接条件变为

$$\mathrm{CAT}(\delta_{i-1}, \delta_i) \Rightarrow \mathrm{CAT}(v^1 - v_i, \delta^1) \wedge \mathrm{CAT}(\delta^1, \delta^2) \wedge \cdots$$
$$\wedge \mathrm{CAT}(\delta^{m-1}, (v_{i+1} - v^m)).$$

一般说来,每条产生式的形式都可以是 N_i 的形式,以使表达式不过分复杂.

由T定义的变换可用$\tau(T)$表示，$\tau(T)$是输入链、输出链的集合:

$$\{(x, y) \mid (S, S) \overset{*}{\Rightarrow} (a(v_1)a(v_2)\cdots a(v_n);\ a(u_1)a(u_2)\cdots a(u_m)\}.$$

可以看出描述标准模式的输入文法是有限状态文法. 由于要表示插入和删除，描述有噪声和畸变的模式的输出文法是上下文无关文法. 已经知道，有限状态属性文法可以作为属性链文法的标准型[4]. 由于插入、删除误差形成了结构发生变化的畸变链，句法部分变得复杂了一些; 又由于在插入m个点时,要求第 1、第m个点分别为 $a(v_i)$, $a(v_{i+1})$ 邻域的点,且m个点之间必须顺次连接,这种连接的要求属于词意的要求(即连接属性),所以把上述变换称为词意、句法指导下的变换.

五、包括词意与句法的距离度量

对于包括小区域的图象识别，若 ω_i 表示第 i 类图象的标准模式，$i = 1, 2, \cdots, K$, 要识别输入模式w属于K类模式中哪一类的方法之一是用统计模式识别方法，选择一组特征,如区域边界的长度、区域的面积、平均灰度等. 如果对第 i 类模式所选择的 n 个特征为 $x_1^i, x_2^i, \cdots, x_n^i$, 则这 n 个特征构成的向量即模式 ω_i 的属性(或词意),

$$A'(\omega_i) = (x_1^i, x_2^i, \cdots, x_n^i)'.$$

定义 2. 一个模式 ω 与 ω_i 之间的属性距离为

$$D_{se}(\omega, \omega_i) = \alpha \| A(\omega) - A(\omega_i) \|$$
$$= \{\alpha_1(x_1 - x_1^i)^2 + \alpha_2(x_2 - x_2^i)^2 + \cdots + \alpha_n(x_n - x_n^i)^2\}^{1/2}.$$

其中 $\alpha_i > 0$, $i = 1, 2, \cdots, n$ 是加权量; $A'(\omega) = (x_1, x_2, \cdots, x_n)'$.

如果特征选择得好,则可以用属性距离最小作为准则进行模式分类,即用典型的统计方法.

在模式发生严重畸变时,如只考虑两个模式之间的属性距离,识别的准确性就会受到影响,因此,考虑有关畸变模式的结构信息是十分重要的. 从标准模式集合到畸变模式集合之间的变换 SSDT 中,删除误差、代换误差及插入误差的数目可以用来表征结构信息. 可以用结构信息定义句法距离如下:

定义 3. 模式 ω 与模式 ω_i 之间的句法距离为

$$D_{sy}(\omega, \omega_i) = \min\{\alpha N_1^i + \beta N_2^i + \gamma N_3^i\}.$$

其中 N_1^i, N_2^i, N_3^i 分别表示用一个 SSDT 把 ω 变换为 ω_i 所需的代换、插入、删除误差的最小的数目；$\alpha > 0$，$\beta > 0$，$\gamma > 0$ 分别表示与代换、插入、删除误差相联系的加权量．实际上 $D_{sy}(\omega, \omega_i)$ 是两个模式之间加权的列维施坦距离．

综上所述，给出既包括词意又包括句法结构的距离度量如下：

定义 4. 模式 ω 与 ω_i 之间包括词意和句法的距离为

$$\begin{aligned} D(\omega, \omega_i) &= D_{se}(\omega, \omega_i) + D_{sy}(\omega, \omega_i) \\ &= \left\{\sum_{j=1}^{n} \alpha_j (x_j - x_j^i)^2\right\}^{1/2} + \min\{\alpha N_1^i + \beta N_2^i + \gamma N_3^i\}. \end{aligned}$$

根据上述距离度量，可以用最小距离准则进行模式识别．由于既考虑了表示统计特征的属性，又加进了表示从标准模式到畸变模式之间结构变化的信息，发扬了句法模式识别方法的长处，从而进一步把统计模式识别与句法模式识别统一了起来．

参　考　文　献

[1] 傅京孙著，模式识别原理及其应用(戴汝为、胡启恒译)，科学出版社，1983.

[2] Tai J. W., Attributed Parallel Tree Grammars and Automata For Syntactic Pattern Recognition. Proc. 5-th ICPR, Miami FL. Dec. 1980.

[3] You K. C. and Fu K. S., A Syntactic Approach to Shape Recognition Using Attributed Grammars, *IEEE Trans* on Vol. SMC-9, No. 6, 1979.

[4] Tai J. W. and Fu K. S., Semantic Syntax-Directed Translation For Pictorial Pattern Recognition, Proc. 6-th ICPR munich Germany, Oct. 1982.

[5] Milgram D. L. and Rosenfeld A., A note on grammar with coordinates, Graphic Languages, Edited by Nake, 1972.

[6] Aho A. V. and Ullman J. D., The Theory of Parsing Translation and Compiling. Vol 1, Parsing Prentice-Hall, 1972.

A SYNTACTIC-SEMANTIC DISTANCE MEASURE FOR PATTERN RECOGNITION

Tai Juwei

(*Institute of Automation, Academia Sinica*)

Abstract

The attributed tree grammar with coordinates for describing noisy and distorted pattern is discussed in this paper. Some semantic rules corresponding to tree productions for describing boundary of a region are given. In addition, based on a semantic syntax-directed translation from pure patterns to distorted patterns, a distance measure including both semantics and syntax is proposed. Thus the minimum distance criterion can be used to recognize the patterns.