

一种基于图匹配的复杂草图识别方法^{*}

张莉莎 孙正兴 周若鸿 徐晓刚

(南京大学计算机科学与技术系 多媒体技术研究所 南京 210093)

摘要 复杂草图识别是手绘草图输入中的一个困难任务。现存的草图识别方法强调的是图形对象简单,但这不适合具有不同复杂性的复杂草图的识别。本文对具有不同复杂性的复杂图形对象提出一个基于图形的统一表示法,文中根据不同信息粒度将复杂图形分别转化为空间关系图(SRG)。文中提出了一个约束的部分枚举,以减小识别复杂草图时匹配 SRG 的状态空间。实验结果显示,我们的方法可适用于具有不同复杂度的各种复杂图形对象的识别。

关键词 草图识别,复杂图形对象,空间关系图,约束的部分枚举

A Method of Graph-Based Composite Sketchy Graphics Recognition

ZHANG Li-Sa SUN Zheng-Xing ZHOU Ruo-Hong XU Xiao-Gang

(Department of Computer Science and Technology, Institute of Multi-media Technology, Nanjing University, Nanjing 210093)

Abstract Composite sketchy graphics recognition is a difficult task in sketch-based graphics input. Existing sketchy graphics recognition ways have emphasized on simple graphic objects and have not suited for recognition of composite graphic objects with different complexity. This paper presents a graph-based uniform representation for composite graphic objects with different complexity, which converts them respectively into Spatial Relation Graph (SRGs) in different information granularity. Accordingly, a Constrained Partial Permutation algorithm is proposed to reduce state space in matching SRGs for recognizing composite sketchy graphics by predicting the directly next state of the current state and neglect invalid states in order. Experiments show that the proposed method can be adapted for recognition of various composite graphic objects with different complexity.

Keywords Sketchy recognition, Composite graphic object, Spatial Relation Graph (SRG), Constrained partial permutation

1 引言

众所周知,正是图形用户界面技术使得计算机应用的普及成为现实。随着计算机及网络技术的不断发展及其向人类工作和生活的不断渗透,对用户界面的要求也越来越高,尤其需要实现以计算机系统为中心向以用户为中心的交互模式转变,自然化、人性化、个性化、智能化成为这一领域的主导,多媒体、多通道支撑下的感知用户界面(Perceptual User Interface)呼之欲出。而手绘草图输入(Freehand Sketch Graphic Input)以模拟现实的纸和笔为目标,它通过将用户的交互意图映射为由压力、方向、位置和旋转等信息组成的多维矢量序列拓宽了人机交互的频带^[1],有利于创造性思想的快速表达、抽象思维的外化和自然交流,从而成为感知用户界面中不可缺少的部分,引起了学术界和产业界越来越多的兴趣。

手绘草图输入的关键技术是以草图图形识别为基础,进而捕捉用户的输入意图。为此,国内外研究者对草图图形识别进行了广泛的研究^[2-3],甚至出现了草图识别用户适应性的研究^[4]。然而,已有的研究主要集中在构成相对简单的图形及其笔划的识别方面^[5-9],其主要存在两个问题:(1)由于所需识别的图形构造比较简单,这些图形一般都由线和弧等所谓基本图元组成。事实上,用户在实际应用中所使用的图形既可能是只包含几个图元的很简单的图形,更大的可能是由种类繁多的多个图元构成的较复杂图形,因此,仅考虑简单图形的

草图表示和识别算法应用在图形结构较复杂的领域内效率不高。(2)为降低识别的复杂度而划分出来的中间逻辑层次往往只针对一些特定复杂度的复杂图形。如:我们前期的研究^[1,2,4,5,9]中将图形识别分解为简单图形识别和复合图形识别两个层次,即在图元之上增加了简单图形(三角形、四边形、五边形和椭圆等),复合图形是由这些简单图形组合而成的。虽然这在一定程度上能降低复合图形识别的复杂度,但简单图形的定义对于不同复杂度的复合图形适应性不强。

针对上述问题,本文将不同复杂层次草图图形元素构成抽象为空间关系图,从而使得不同复杂层次的草图具有统一的表示,而将不同复杂程度的图形识别转化为不同信息粒度的空间关系图相似度计算,并在空间关系图相似度计算过程中动态生成匹配约束条件,根据生成的匹配约束条件使用约束的部分枚举算法智能预测当前匹配状态的后继有效的枚举状态以跳过无效的中间匹配状态,达到状态空间剪枝的目的,有效地降低了空间关系图匹配过程中状态搜索空间,从而可以适应不同复杂程度的图形识别,且其识别结果直接为草图语义的理解奠定了基础。

2 复杂图形的描述及属性关系表示

2.1 复杂图形及其构成

复杂图形通常是由多个线、弧和圆等图元按一定几何关系构成的。实际应用中,这种复杂图形都可以看成是由若干个

^{*} 本文得到国家自然科学基金项目(编号:69903006、60373065)资助。张莉莎 硕士生,研究方向为智能用户接口;孙正兴 副教授,博士后,研究方向为智能用户接口、多媒体挖掘和图形图像技术;周若鸿 硕士生,研究方向为智能化人机界面;徐晓刚 硕士,研究方向为智能用户接口。

结构相对简单的图形(以下称为图形元素)通过某种空间关系模式构成的。图形元素的定义和选择与应用领域密切相关,且具有递归性,这可采取基于特征设计中的形状特征定义及其装配建模方式^[10~12]来确立。为有效描述复杂图形,我们将复杂图形看成是一个类对象,每个类对象都是由属性信息及构成这个对象的图形元素间的空间关系信息来描述。其中:属性信息包括这些图形元素的类型、坐标和尺寸等,空间关系信息包括图形元素之间的相对空间位置、相对方位和相对旋转角度等。这样,复杂图形描述相应转化为不同信息粒度之上的属性和空间关系表示,以适应不同复杂程度的图形描述。

当把用户手绘的复杂图形与目标复杂图形比较时,图形元素的属性信息比较相对简单,因而识别重点在于图形元素

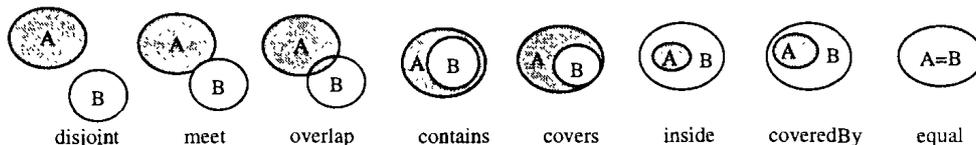


图1 八种典型的拓扑结构关系

我们使用上述的八种拓扑关系来描述复杂图形中的图形元素的拓扑约束,由于 covers 关系和 contains 关系、covered by 关系和 inside 关系之间没有非常大的差异,因此使用这两种关系进行空间拓扑关系描述时,无需严格区分,可以简化成六种拓扑关系。

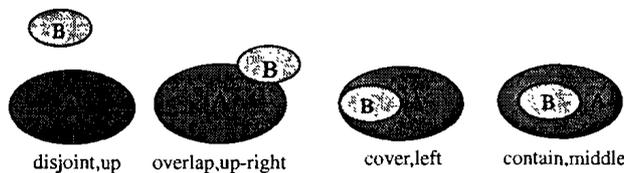


图2 方位约束示例

2.1.2 方位约束描述方法 图形元素的拓扑约束描述

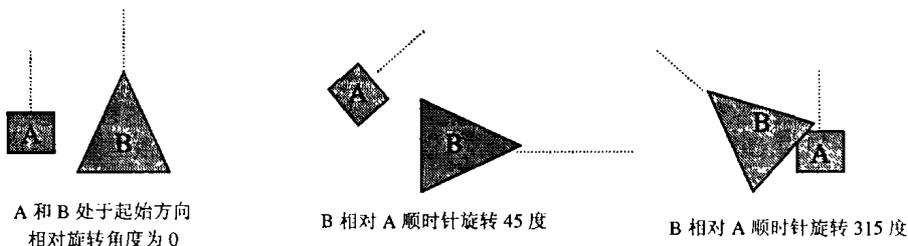


图3 旋转约束示例

上述空间关系模式的三层约束可以保证复杂图形对象的属性和结构是唯一的,且图形具有平移不变性、缩放不变性和旋转不变性。

2.2 复杂图形的 SRG 属性关系描述

空间关系图(Spatial Relation Graph, SRG)是一种把统计和结构方法结合起来描述模式的结构。其结点形象地表示一个复杂图形的图形组成元素,边表示各个图形元素之间的关系,同时又给它的结点和边赋予相应的属性值,来描述元素结构之外的模式特征。由此,使用空间关系图可以形象地描述复杂图形对象模式的所有信息,包括结构信息和结构之外的统计模式信息。所有针对模式之间的计算也相应地转换成为对应的空间关系图之间的计算。下面我们定义空间关系图。

定义1(空间关系图) 一个空间关系图可以表示为一个

之间的空间关系模式。空间关系模式包括拓扑约束,方位约束和旋转约束。拓扑约束是指复杂图形内部的构造信息,如:相交、包含等,方位约束是指图形元素之间相对位置的方向性,如上、下、左、右等;旋转约束是指图形元素之间的相对旋转角度。

2.1.1 拓扑约束描述方法 描述拓扑关系的方法很多,其中比较著名的是 9-Intersection 方法^[13]。9-Intersection 是描述点、线、面的空间拓扑关系的定性形式化模型,它的理论基础是集合拓扑,但是它的 512 种关系中有一些是根本无法实现的^[13]。我们对两个二维平面的无洞的区域 A 和 B 归纳出八种典型的空间拓扑关系,如图 1 所示。

只表示图形对象的凸包具有某种邻接关系,我们需要更进一步地表示邻接在什么位置,也就是描述方位约束。方位是指一个图形元素相对于另一个图形元素的位置的方向,由两者凸包的中心连线的斜率决定,一般分为九个方向:上,下,左,右,左上,左下,右上,右下和居中,图 2 示出部分 B 相对于 A 的方位约束例子。

2.1.3 旋转约束描述方法 旋转约束是各个图形元素相对于坐标轴的旋转角度之间的约束关系,一般可以将圆周等分成八个旋转区域。用下述空间关系图描述时,图形元素的旋转角度可以记录在该元素对应的结点信息中,而两两之间的相对旋转角度可以记录在对应的边信息中。如图 3 的示例中,虚线表示图形元素的起始方向,其他旋转方向都是顺时针相对于起始方向的旋转角度。

六元组 $G=(V, E, A_V, \mu, A_E, \gamma)$, 其中:(1) V 是空间关系图的结点集,即: $V=\{V_1, V_2, \dots, V_p\}$, p 为结点的个数;(2) $E: V \times V$ 是空间关系图的边集,即: $E=\{V_1V_1, V_1V_2, \dots, V_pV_p\}$;(3) A_V :是空间关系图的结点的属性集的集合,即: $A_V=\{A_1^V, A_2^V, \dots, A_p^V\}$, A_i^V 是第 i 个结点的属性集;(4) A_E :是空间关系图的边的属性集的集合,即: $A_E=\{A_1^E, A_2^E, \dots, A_p^E\}$, A_i^E 是第 i 条边的属性集;(5) $\mu: V \rightarrow A_V$, 是生成每个结点属性的函数,它为每个结点分配一个对该结点有意义的属性;(6) $\gamma: E \rightarrow A_E$, 是生成每个边属性的函数,它为每条边分配一个对该边有意义的属性。

3 基于 SRG 的图形识别和匹配算法

3.1 基于 SRG 的匹配策略

在过去的几十年中,研究者们提出了大量的图匹配算法来降低图匹配计算的复杂性^[14,15]。本文使用 SRG 描述复杂图形的模式信息,那么基于 SRG 的信息表示就是使用 SRG 描述复杂图形的三层约束。组成复杂图形的图形元素被抽象成 SRG 的结点,把图形元素之间的拓扑约束和方位约束抽象成为结点之间的边,结点的属性信息主要包含图形元素的类型,凸包的位置、大小和图形元素相对坐标轴的旋转角度等,边的属性信息主要包含拓扑关系类型和方位类型。

当我们使用 SRGs 来表示复杂图形的整体信息后,复杂图形之间的匹配就转化为计算它们对应的 SRGs 的相似度。

定义 2(空间关系图的匹配) 空间关系图的匹配是一个从源空间关系图 $G(V, E, A_v, \mu, A_e, \gamma)$ 到目标空间关系图 $G'(V', E', A_{v'}, \mu', A_{e'}, \gamma')$ 的单映射 $\eta: V \rightarrow V'$, 该映射满足: (1) $\forall v \in V, \mu(v) = \mu'(\eta(v))$; (2) $\forall e = (v_i, v_j) \in E, \exists e' = (\eta(v_i), \eta(v_j)) \in E' \wedge \gamma(e) = \gamma(e')$ 。

从上面的定义可以看出空间关系图的匹配包含两层含义:属性匹配和结构匹配。我们将使用空间关系图实时地表示用户绘制的复杂图形来进行模式匹配的计算。若直接从定义 2 实现空间关系图的匹配复杂度会很高,因而我们采用文[2]中计算图匹配的优化算法——约束的部分枚举算法。

3.2 约束的部分枚举算法

枚举定义 2 中的映射 η 是一个全排列问题,而全排列问题的算法复杂度是 P_n^m (n 是目标空间关系图中结点的个数, m 是源空间关系图中结点的个数)。然而,空间关系图的匹配至少要满足结点匹配和边匹配这两个约束条件,因此我们可以根据定义 3 中描述的几个约束条件,在枚举过程中不断产生顶点匹配约束和边匹配约束。

定义 3(约束的部分枚举) 给定两个正整数 m 和 $n, m \leq n$, 从 $[1..n]$ 选取 m 个不同的整数把它们排成一个序列,记为 $B_1 B_2 \dots B_k \dots B_m$, 其中 $1, 2, \dots, k, \dots, m$ 是这个整数表中的位置, $B_1, B_2, \dots, B_k, \dots, B_m$ 是这个整数表中相应位置的值,计算所有至少满足下面两个约束条件之一的序列的过程被称为约束的部分枚举: (1) 顶点匹配约束: i 不能放在位置 k , 即 $B_k \neq i$, 记为 $R_v(i, k)$; (2) 边匹配约束: 如果 i 已经放在位置 k , 则 j 决不能放在位置 l , 即如果 $B_k = i$, 则 $B_l \neq j$, 其中, $i, j \in [1 \dots n]$; $k, l \in [1 \dots m]$ 并且 $k < l$, 记为 $R_e(i, k) - (j, l)$ 。这实际上是一个有条件顶点匹配约束。

我们根据这些匹配约束条件,动态地调整增长点 d 的位置,使算法从当前序列状态略去部分无效的序列状态跳到它的直接后继状态。

定义 4(Next_d-Permutation (S) (1 ≤ d ≤ m)) 给定整数序列 $S = S_1 S_2 \dots S_d \dots S_m$, S 的位置 d 的直接后继状态 Next_d-Permutation 是 $C = C_1 C_2 \dots C_m$, 其中, $S_1 S_2 \dots S_d < C_1 C_2 \dots C_d$, 并且不存在另外一个整数序列 $E = E_1 E_2 \dots E_m$ 满足 $S_1 S_2 \dots S_d < E_1 E_2 \dots E_d < C_1 C_2 \dots C_d, C_{d+1} C_{d+2} \dots C_m$ 是 $\{1, 2, 3, \dots, n\} - \{C_1, C_2, \dots, C_d\}$ 这 $(n-d)$ 个数中最小的 $(m-d)$ 个数的从小到大的排列。我们称 d 为增长点。

例如,如果当前整数序列 $S = 12345$, 增长点 $d = 3$, 则 S 的位置 d 的直接后继状态是 12435, 而不是其他序列,如 12354, 12534, 12543 等等。

因此,在全排列过程中许多无效的枚举状态被直接略去不予考察,节省了大量的计算。

4 算法复杂度分析和实验结果

在匹配过程中,当某个顶点匹配约束条件 $R_v(i, k)$ 满足

时,计算直接后继状态时略过的枚举状态的个数为 $(P_{(n-k)}^{(n-k)} - 1)$, 在整个枚举过程中,我们使用该顶点匹配约束条件 $R_v(i, k)$ 进行后继状态计算总共 $P_{(n-k)}^{(n-k)}$ 次。因此,使用该顶点匹配约束条件 $R_v(i, k)$ 而直接略去的中间枚举状态的个数总共为 $P_{(n-k)}^{(n-k)} * (P_{(n-k)}^{(n-k)} - 1) = P_{(n-k)}^{(n-k)} - P_{(n-k)}^{(n-k)}$, 这是一个关于三个参数 n, m 和 k 的函数,我们把使用该顶点匹配约束条件而略去的中间枚举状态集合记为 $Vertex(R_v)$, 该集合的元素个数为 $P_{(n-k)}^{(n-k)} - P_{(n-k)}^{(n-k)}$ 。同理,使用某个边匹配约束条件 $R_e(i, k) - (j, l)$ 而直接略去的中间枚举状态的个数总共为 $P_{(n-2)}^{(n-2)} - P_{(n-2)}^{(n-2)}$ 。这是一个关于三个参数 n, m 和 l 的函数,我们把使用该边匹配约束条件而略去的中间枚举状态集合记为 $Edge(R_e)$, 该集合的元素个数为 $P_{(n-2)}^{(n-2)} - P_{(n-2)}^{(n-2)}$ 。

假设,总共有 q 个顶点匹配约束条件和 r 个边匹配约束条件,根据这些匹配约束条件直接略去的中间枚举状态是分别使用各个匹配约束条件而直接略去的中间枚举状态集合 $Vertex(R_{v1}), Vertex(R_{v2}), \dots, Vertex(R_{vq}), Edge(R_{e1}), Edge(R_{e2}), \dots, Edge(R_{er})$ 的并集,我们把它记为 US 。由于状态集合中略去的枚举状态有交集,计算集合 US 的大小是很复杂的问题,我们可从实验上直观地考察。

在实验中,预先创建了 300 个具有代表性的复杂图形^[2], 这些图形对象可以看成由预先定义的、一般不超过 6 个图元、结构相对简单的图形元素构成。把用户绘制的每个查询图形对象跟所有复杂图形进行相似度计算。主要考察算法的平均优化率,它是在匹配过程中剪枝剪掉的无效中间枚举状态的个数占所有枚举状态总数的比率,这个指标能够比较有力地说明算法应用在复杂图形识别的性能。平均优化率随着 m, n 变化的曲线如图 4 所示。我们还使用约束条件覆盖率这个参数来衡量匹配过程中生成的匹配约束条件的个数占所有可能的匹配约束条件总数的比率,它反映的是生成的匹配约束条件的整体情况,而不关心具体每条匹配约束条件所在的位置。我们分别使用公式(1)(2)(3)来计算特定的 m, n 和约束条件覆盖率(Constraint Coverage, cc)所对应的平均优化率。

$$AOR(m) = \left(\sum_{i=1}^M \left(\sum_{j=1}^N \frac{optimized(m, n_j)}{N} \right) \right) / M \quad (1)$$

$$AOR(n) = \left(\sum_{j=1}^N \left(\sum_{i=1}^M \frac{optimized(m_i, n_j)}{M} \right) \right) / N \quad (2)$$

$$AOR(cc) = \sum_{i=1}^K \frac{optimized(query_i)}{K} \quad (3)$$

其中, AOR 是平均优化率的英文缩写, m 是查询图形对象的构成元素的个数, M 是拥有 m 个构成元素的图形对象的个数, N 是拥有不少于查询图形对象的构成元素个数的构成元素的复杂图形的个数, $optimized(m, n_j)$ 是查询图形对象跟图形库中某个特定的复杂图形匹配时的优化率, m_i 是拥有 m 个构成元素的查询图形对象的索引号, n_j 是被匹配的图形库中的复杂图形的索引号, 该复杂图形拥有 n_j 个构成元素, 并且 $n_j \geq m_i$ 。 K 是约束条件覆盖率为 cc 的查询图形对象的个数。

图 4 显示的平均优化率是非常高的。图 4(a) 中平均优化率随着 m 变化的曲线是我们提交查询图形对象时需要特别关注的, 当复杂图形库相对稳定时, 匹配计算的效率很大程度上完全决定于该图形对象本身。当 m 大于 10 时, 平均优化率的值大于 0.9999319119, 即优化以后的计算量还不到原始计算量的万分之一。当 m 大于 4 时, 平均优化率也达到 95% 以上。平均优化率随着 n 变化的值也非常高。从图 4(a) 和图 4(b) 中, 我们可以观察到平均优化率随着 n 变化的曲线的曲率

比它随着 m 变化的曲线的曲率小,这是因为我们使用了不同的计算方法。根据查询图形对象和目标复杂图形间的部分匹配的定义,我们可以发现 $m \leq n$, 计算对 m 的平均优化率是所有不少于 m 个构成元素的复杂图形的平均,而对 n 的平均优化率是所有不多于 n 个构成元素的查询图形对象的平均。图

4(c)中,平均优化率随着约束条件覆盖率波动比较大,这说明不单是匹配约束条件的覆盖率影响平均优化率,匹配约束条件的分布更多地影响平均优化率。例如很多匹配约束条件对应的位置是 m 的话,它们对于平均优化率的贡献是微乎其微的,因此大的约束条件覆盖率可能对应到很小的平均优化率。

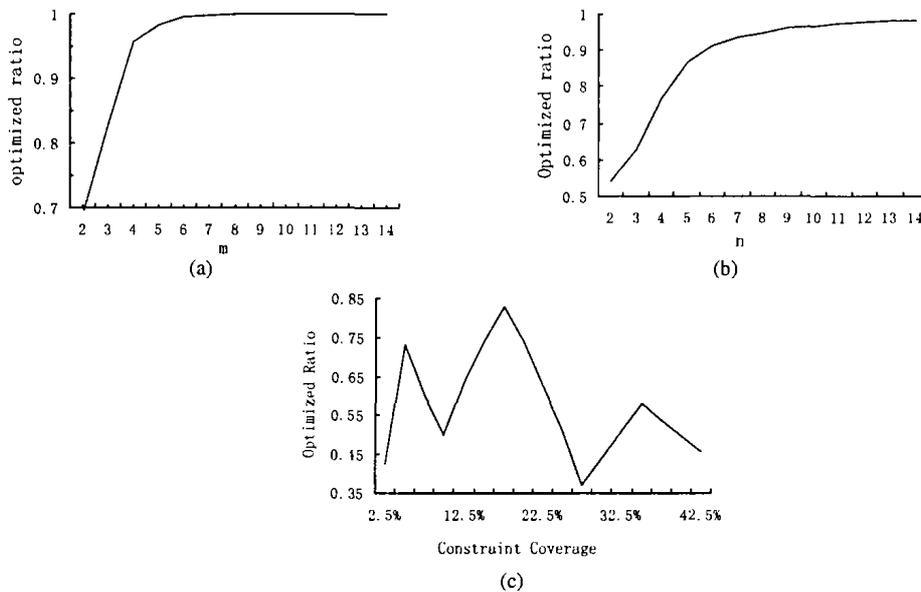


图 4 平均优化率随 m 、 n 和 cc 变化的曲线

总结与展望 手绘草图输入的关键技术是以草图图形识别为基础,进而捕捉用户的输入意图。已有草图图形识别系统的识别对象主要是构造比较简单的图形,或者在面对复杂的图形结构时划分出一个简单图形层次,很难适应不同复杂度的复杂图形识别。针对这些问题,本文将结构相对复杂的图形视为由结构相对简单的图形元素构成,从而使得不同复杂度的草图图形具有统一的表示,并将不同复杂度的复杂图形识别转化为不同信息粒度的空间关系图相似性计算,而后在空间关系图相似度计算过程中动态生成匹配约束条件,并根据生成的匹配约束条件使用约束的部分枚举算法智能预测当前匹配状态的后继有效的枚举状态以跳过无效的中间匹配状态,达到状态空间剪枝的目的。实验表明当复杂图形包含 4 个以上的图形元素时,匹配算法的平均优化率达到 95% 以上,有效地减少了空间关系图匹配过程中的状态搜索空间,从而能快速识别结构较为复杂的图形。由于不同复杂程度图形元素间的统一描述,SRG 的表示和匹配算法适合于不同复杂度的草图图形识别。草图图形的识别结果直接为我们进一步针对特定的应用领域进行草图语义的解释奠定了基础。

参考文献

- 1 孙正兴,徐晓刚,孙建勇,金翔宇. 支持方案设计的手绘图形输入工具. 计算机辅助设计与图形学学报,2003,已录用
- 2 徐晓刚. 草图理解系统及其关键技术研究.[南京大学计算机科学与技术系硕士学位论文]. 2003. 5
- 3 Hearst M A, Gross M D, Landay J A, Stahovich T F. Sketching Intelligent Systems. IEEE Intelligent Systems, 1998, 13(3): 10~19
- 4 彭彬彬,金翔宇,徐晓刚,孙正兴. 基于 SVM 增量学习的用户适应性研究. 计算机科学,2003(3)

- 5 孙建勇,金翔宇,彭彬彬,孙正兴. 一种快速在线图形识别与规整化方法. 计算机科学,2003(2)
- 6 Lank E, Thorley J S, Chen S J-S. An interactive system for recognizing hand drawn UML diagrams. In: Proc. of CASCON, 2000
- 7 Hammond T, Davis R. Tahuti: a geometrical sketch recognition system for UML class diagrams. AAAI Spring Symposium on Sketch Understanding, California, 2002. 59~68
- 8 Calhoun C, Stahovich TF, Kurtoglu T, Kara LB. Recognizing Multi-Stroke Symbols. AAAI Spring Symposium on Sketch Understanding, California, 2002
- 9 Liu W Y, Jin X Y, Sun Z X. Sketch-Based User Interface for Inputting Graphic Objects on Small Screen Device. Lecture Notes in Computer Science, Springer, 2002, 2390: 67~85
- 10 孙正兴,张福炎. 方案设计中的特征定义与表示方法. 中国机械工程, 2001, 10(10): 1182~1185
- 11 Sun Z X, Zhang F Y. A formalized approach for describing feature-based design methodology. J. of Computer. Science & Technology, 1998, 13(Supplement): 128~130
- 12 孙正兴,张福炎,蔡士杰. 一种面向方案设计的装配建模表示方法. 计算机辅助设计与图形学学报, 2000, 12(1): 28~33
- 13 Egenhofer M, Franzosa R. Point-Set Topological Spatial Relations. International Journal of Geographical Information Systems, 1991, 5(2): 161~174
- 14 Foggia P, Sansone C, Vento M. A Database of Graphs for Isomorphism and Sub-Graph Isomorphism Benchmarking. In: Proc. of 3rd IAPR-TC15 Workshop on Graph-Based Representations in Pattern Recognition, Ischia. 2001
- 15 Messmer BT, Bunke H. Efficient Sub-graph Isomorphism Detection: A Decomposition Approach. IEEE transactions on knowledge and data engineering, 2002, 12(2): 307~323