

在线草图识别中用户手绘习惯建模方法^{*}

张 斌 孙正兴 孙建勇 彭彬彬

(南京大学计算机软件新技术国家重点实验室 计算机科学与技术系 南京210093)

摘 要 手绘草图是概念设计和思路外化的一种高效的表达方式。用户绘制草图时存在的多种形式,及其随意性和模糊性使得用户适应性问题逐渐成为草图识别的核心课题。本文提出了一种在线草图识别的用户建模方法来捕捉绘制草图时的用户习惯,主要包括两个方面的内容:一是基于SVM的主动式增量学习方法,二是基于动态用户建模的手绘复杂图形的识别方法。前者与传统的增量式学习方法相比,在识别精度相同的情况下所需的训练时间和训练数据集要少得多。后者则是基于笔划信息以及笔划间的顺序和空间关系信息,采用增量式决策树捕捉用户的输入习惯和过程信息。实验证明了本文方法在在线草图识别中的有效性和高效性。

关键词 在线手绘图形识别,用户适应性,用户建模,增量式主动学习,支撑向量机(SVM),模糊预测

User Modeling for on-Line Sketchy Graphic Recognition

ZHANG Bin SUN Zheng-Xing SUN Jian-Yong PENG Bin-Bin

(National Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

(Computer Science and Technology Department, Nanjing University, Nanjing 210093)

Abstract Freehand sketching is an efficient way for idea externalizing and concept design. User adaptation to be the center of the one of sketch recognition, because of multiformity, arbitrary and fuzziness during users' sketching processes. This paper presents a user modeling method of online sketch recognition to capture users' sketching habits, including two aspects; Strokes recognition based on SVM-based active incremental learning and composite sketchy graphics recognition based on dynamic user model. The former, compared to the normal incremental learning method, needs much less learning time and small training dataset under the same recognition precision. The latter, based on the stroke's information and the special relation and sequence of strokes, adopts the incremental decision tree to capture users' input habits and process information. Experiments show it both effective and efficient in online sketchy graphic recognition.

Keywords On-line sketchy graphics recognition, User adaptation, User modeling, Incremental active learning, Support vector machines (SVM), Fuzzy prediction

1. 引言

手绘草图是人类一种自然而直接的思路外化和概念设计的方式^[1]。草图识别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笔划矢量序列,同时,根据人对形状的视觉认知规则、思维习惯和信息模型对用户输入草图及其过程进行猜测和推理,进而捕捉和理解用户输入意图。然而相同的输入意图可能会因用户的不同而有不同的草图表现,不同用户存在着不同的手绘习惯和特点,因此用户适应性研究在草图识别中对识别效果起着极为关键的作用。尽管草图识别技术的历史可追溯到 Ivan Sutherland 的 Sketchpad 系统^[2],20世纪70年代就已出现了对草图识别的研究^[3,4],尤其是20世纪90年代小型化、便携式和无线设备大量涌现及“平板计算机(Tablet Computer)”对笔式交互应用的需求极大地推动了这一领域的发展,已经出现了大量的研究和应用成果^[5]。然而,由于草图涉及的范围更大,样本收集比较困难,而且其生成具有更大的随意性和自由度,因此,与说话者识别^[6]或书写者识别^[7]相比,草图识别的用户适应性具有更大的难度,相关研究也较少见。

草图识别过程主要包括以下步骤^[8,9]:输入的笔划在预处

理之后抽取笔划特征,并对笔划进行识别,识别后的笔划成为草图图形识别的基础;而后,进行草图图形识别。我们在前期研发的原型系统^[8,9]中,分别采用 Rubine^[10]的统计方法进行笔划特征抽取;基于模板匹配方法进行笔划识别;基于规则

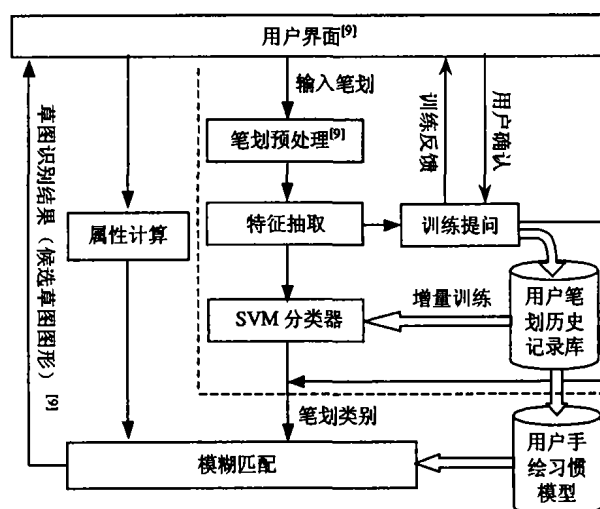


图1 面向用户适应性的用户建模框架

^{*} 本文得到国家自然科学基金项目(6990300、60373065)资助。张 斌 硕士生,主要研究方向为在线图形识别;孙正兴 博士,副教授,研究方向为视觉信息挖掘、智能用户接口、多媒体辅助工程;孙建勇 硕士生,主要研究方向为在线图形识别;彭彬彬 硕士,主要研究方向为机器学习、智能化人际接口。

的方法进行草图图形识别,无法满足用户适应性的要求。为此,本文设计了如图1所示的面向用户适应性的草图识别方案。该方案要有两个特点:一是采用转角函数进行笔划特征抽取和基于支撑向量机(SVM)主动式增量学习方法进行笔划识别;二是采用增量决策树和模糊匹配的方法来构造用户手绘习惯模型和对用户手绘草图图形进行识别。本文将重点介绍这两个主要技术,而其它过程见文[8,9]。

2. 手绘草图笔划识别

2.1 笔划的特征抽取

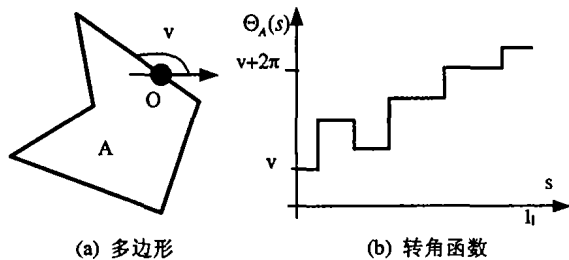


图2 多边形与转角函数

目前,在草图识别领域,笔划的几何特征的选择都充分考虑了平移、旋转和缩放不变性。Esther^[12]提出的转角函数方法,由于其在多边形识别上的特殊优势,以及我们的研究背景—UML 概念设计—不需要强调旋转不变性这两个原因,被我们选为待抽取的笔划特征。转角函数 $\Theta_A(s)$ 是指图形 A 上关于弧长 s 的一个函数,它标识了从图形 A 边界上的初始点 O 开始到当前点(沿逆时针方向转过弧长为 s 的点)累积的转角(逆时针方向为正,顺时针方向为负)。其中 $\Theta_A(o)$ 为选取的初始点 O 的切线方向与某一特定方向(例如水平方向)的夹角,如图2所示。为了简化问题和更好地适应 UML 概念设计,在预处理阶段我们对转角函数的应用加入了一定的限制条件,并对记录下的用户输入草图进行了规一化处理,详见文[8,9,11]。改进后的转角函数定义如下:

$$V_i = \Theta_A\left(\frac{i \bmod d}{d}\right) - \Theta_A\left(\frac{i-1}{d}\right) \quad 0 \leq i \leq d$$

其中: d 为需要抽取特征的维度,即将图形的边界 d 等分。

上述方法所抽取的是单笔划特征。对于多笔划复杂草图,我们主要关心的是如何有效地抽取笔划的顺序、几何构成和连笔特征。为此,我们按顺序将用户所绘草图中每一笔的落点和下一笔的起点用线段连接,使整个草图连成了一条连续的曲线(如图3所示)。经过这样的处理之后,所得的信息其实包括两个部分:一部分是笔划,即用户所画的部分(图3中黑线);另一部分是关系,即系统根据上述规则加上的部分(图3中浅色线条)。将此连续曲线看作一个整体,就把这个多笔划草图转换成了一个类似的单笔划草图,可以采用单笔划的特征抽取方法对其进行特征抽取。笔划的连接由于是遵循笔划的输入顺序的,因此能有效地表示笔划的顺序特征。这种方法对复杂图形的抽取具有一定的通用性。



图3

2.2 SVM 主动式增量学习方法

通常,用户总是一笔一笔地输入草图图形,然后由系统收集用户绘制的笔划信息,系统利用这些收集到的数据来训练 SVM 的分类器,尽管这些经过训练得到的 SVM 分类器对用户具有很强的适应性,但仍然存在两个问题:第一,当系统随着用户样本数据的积累,分类器的性能虽然不断地增长,但是如果全部重新训练这些分类器,那么时间的增长也是非常迅速的,如何能有效地利用 SVM 的历史训练数据,大大缩短训练的时间呢?第二,训练 SVM 分类器需要使用人工标注过的样本,即必须要有这些笔划的类别信息;如何能有效减少人工标注的工作量也是一个很有意义的问题。为解决第一个问题引入了 SVM 增量学习方法^[13,14],取得了较好的效果。对于第二个问题,我们在增量式学习的基础上,引入主动学习的概念,提出了主动式增量学习的方法。

在训练分类器过程中,对分类器有作用的是那些少数系数 α_i 不为零的支持向量(SV),其余大部分样本都没有作用。因此,在收集数据过程中,仅仅收集那些对分类器训练有用的样本就可以了,虽然我们不可能根据未知的数据获得未知的知识,但可使收集的数据集合($S_{Collected}$)尽量地小,仅需包含这个 SV 集合(S_{SV})就行了,即: $\text{Min}(S_{Collected} | S_{SV} \subset S_{Collected})$ 。所谓主动学习,就是系统根据已经训练的结果来对用户输入的笔划信息进行一次评价,把那些对分类器训练有用的样本拿出来向用户提问(称为主动提问),用户仅仅需要标注系统提问的样本数据的类别,这些样本在整个输入样本中只占一个很小的比率,这将极大地减少人的工作强度。

显然,对用户样本的重要程度的评价标准是关键所在。在训练过程中, α_i 不为零的样本被选为支持向量,这些样本在映射后的空间上表现为靠近超平面。利用这个性质,我们可以根据样本点在映射后跟超平面的距离来决定该样本点的重要性,若样本点离超平面足够近,就向用户提问其类别属性,并把它选入 $S_{Collected}$;否则,可以忽略该样本点。

若记初始训练样本集为(IS),增量样本集为(INS),临时样本集为(NS)和工作集为(WS),那么增量式主动学习算法可以描述为:

- Step 1. $\Gamma = \text{Train}(IS), WS = IS_{SV}$;
- Step 2. $NS = \text{Query}(INS), WS = WS \cup NS$;
- Step 3. $\Gamma = \text{Train}(WS), WS = WS_{SV}$;
- Step 4. goto step 2.

对于多值情况,只需对每个分类器做一个提问,然后挑选出每个分类器中重要的数据即可。

3. 用户手绘习惯建模

用户绘制习惯和方式的不同对有效地识别草图和推测用户意图带来了很大困难。尽管基于 SVM 的草图识别方法对于笔划识别可以达到较好的识别效果,但是其应用于多笔划复杂草图识别的效果却并不令人满意。为此,我们提出了利用用户输入笔划的类别、笔划的方向、笔划间的空间关系及顺序等笔划属性信息来建立用户手绘习惯模型的方法。笔划的类别是通过增量式主动学习方法进行识别或者通过用户的反馈而获取的;笔划的方向及笔划间的关系是通过基于规则的方法计算得到^[9];而用户习惯模型是采用决策树方式、按照用户输入笔划的顺序建立的;当用户输入一个草图对象时,就可以从决策树的根节点开始,以用户输入的笔划为单位,逐次向下一层查找匹配的结点,通过模糊匹配,预测用户的输入意图,识别用户输入的草图对象。

3.1 手绘习惯模型的表示和建立

在构造用户模型的过程中,我们主要使用三类信息,即用户的笔划信息、笔划间关系信息和笔划顺序信息,分别定义如下:

笔划信息 s 包括笔划的类别信息 id 和方向信息 dir 。在训练过程中,由用户进行指导而得到笔划的准确类别;在分类过程中,通过抽取输入单笔划的特征,然后利用 SVM 分类器来识别其类别信息,由于识别精度并不是100%,从而需要进行模糊校验。 dir 为笔划的方向信息。由于用转角函数的方法抽取出来的特征值不满足旋转不变性,因此需要记录笔划的方向信息。我们采用坐标八分限的八个方向来标识笔划的方向,即:将 x 轴作为正方向也就是 I 方向,将其逆时针旋转 $\pi/4$ 得到其余七个方向,分别为 II、III、...、VIII,从而一共得到8个方向;对于封闭的笔划,使用0来标记它的方向。

笔划间的关系信息 r 主要记录笔划与其前一笔划间的相对位置关系信息,我们首先计算两个图形的重心,然后从前一笔划向当前笔划作一矢量线,按其方向分成上、下、左、右、左上、左下、右上、右下8个方向,分别用 I、II、...、VIII 标识。对于

每个图形的第一个笔划不存在关系信息,则用0标识。

笔划间的顺序信息由用户绘画笔划的相对先后顺序自然得到,反映在用户手绘模型中,就是某个用户模型树中从根结点到叶节点的一条路径。

考虑了建立用户模型所需的三类信息以后,我们进一步考虑用户模型的表示方式。考虑到效率和实现的方便,我们采用决策树学习的思想来表示用户模型。

为了降低用户建模的搜索策略的复杂度,我们存储每个节点向下搜索到的可能图形类别和权重,定义其为预测列表 $GList$,表中每一项包含两个信息: Gra_id 为可能图形的类别, $weight$ 为该图形的权重(表项的顺序按该权重从大到小排列)。

单个用户的模型可以表示为树 $T=(V, E)$ 。 V 为树的节点信息集合,即: $V=(v_1, v_2, \dots, v_l)$, $v_i=(s, r, GList)$, $s, r, GList$ 分别是上述的笔划信息,关系信息和预测列表。 E 为树的边集,记录笔划间的顺序信息。多棵用户模型树合在一起成为一个森林,这就是用户手绘习惯模型。用户模型树可以动态添加和调整,具体调整过程见文[15]。

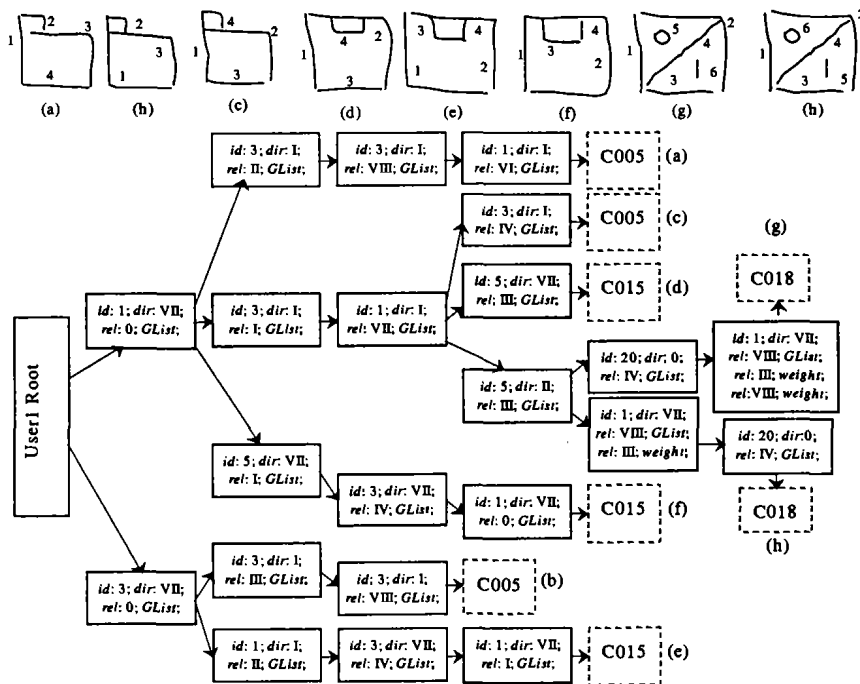


图4 草图图形的画法及其模型表示示例

图4示出了我们草图图形库中标号为 C005、C015 和 C018 图形的可能画法及其用户手绘模型。

3.2 草图图形的模糊预测

手绘草图固有的模糊性需要用用户模型检索提供相应的一定模糊机制。为此,我们在草图识别和检索时采取了同时监控树中同一层多个结点的方式,这保证了用户建模提供在线用户帮助的可用性,而且在实验中也得到了很好的效果。所处理的模糊性主要包括两类:

① 方向的模糊性:例如:一个笔划的方向与 x 轴成 25° 角,按照属性的判断规则,那么其方向应该是 II,但是,用户可能真实的绘制方向为 I,因此,笔划的方向信息必须满足一定的模糊性。我们定义方向的隶属函数为:

$$f(x) = \begin{cases} \frac{1}{2} + \frac{1}{2} \sin(4x - \frac{2i-3}{2}\pi) & x \in (\frac{(i-2)\pi}{4}, \frac{i\pi}{4}) \\ 0 & x \leq \frac{(i-2)\pi}{4} \quad x \geq \frac{i\pi}{4} \end{cases}$$

其中: $i=1, 2, \dots, 8$, 分别对应 I, II, ... , VIII 等8个方向, x 为弧度。

② 笔划类别的模糊性。笔划类别的模糊性主要是由 SVM 分类器的分类错误引起的。我们根据对 SVM 分类器测试的结果,定义了模糊映射矩阵 R ,如图5所示,图中:横向 j 表示分类器分类的类别,纵向 i 为笔划的实际类别,矩阵内的数据表示分类器分类类别为 i 而实际类别为 j 的权重(空为0)。

令 g' 是当前某监控结点的权重值, $GList$ 为该结点的预测列表,则列表中各候选对象的权重值为 $g' \cdot GList$ 。当用户新输入一笔时,该笔划模糊匹配得到的方向记为 s_1 和 s_2 ,与前一笔划的关系记为 r_1 和 r_2 ,可能的图形类别为 R_n 。则,该笔划所对应的搜索结点的权重为: $g = g' s_m r_n R_n$,其中 $m=1$ 或 $2, n=1$ 或 $2, i, j=1, \dots, 14$ 。通过这样的一个递归过程,我们就完成了权重的调整 and 用户模型的模糊检索。

$$R = \begin{matrix} i \setminus j & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \end{matrix} & \begin{matrix} 0.7 & 0.2 & & 0.05 & 0.05 & & & & & & & & & & \\ 0.1 & 0.9 & & & & & & & & & & & & & \\ 0.05 & 0.05 & 0.8 & & & & & & & & 0.1 & & & & \\ 0.1 & 0.1 & & 0.6 & & 0.05 & & & & 0.05 & 0.05 & & & 0.05 \\ & & 0.1 & & 0.05 & 0.85 & & & & & & & & & \\ & & 0.1 & & & 0.05 & 0.85 & & & & & & & & \\ & & & & & & & 0.95 & 0.05 & & & & & & \\ & & & & & & & & 1 & & & & & & \\ & & & & & & & & & 1 & & & & & \\ 0.05 & 0.05 & 0.05 & & & & & & & & 0.85 & & & & \\ 0.05 & & & 0.05 & & 0.05 & 0.05 & & 0.05 & & 0.55 & 0.2 & & & \\ & & & & & & 0.05 & & & & & & 0.85 & & 0.1 \\ & & & & & & & & & & & & & 1 & \\ & & & & & & & & & & & & & 0.1 & 0.9 \end{matrix} \end{matrix}$$

图5 笔划类别的模糊映射矩阵

4. 实验与分析

4.1 实验样本的收集

我们分析了 UML 设计和电路图设计中的常用图符,从图形组成的角度选取了150个复杂图形来进行草图识别算法的评测和对比。在样本采集中,我们共收集了以上150类复杂图形不同画法的5370个数据,其中每种画法收集了10个样本,如图6所示,这些收集的数据部分作为用户建模的训练数据,另一部分作为用户建模的测试数据。同时,从这150类的复杂图形中,我们抽取出了14种用户常用的笔划,如图7所示并收集了这14类笔划的76452个记录,随机抽取其中53200个记录作为训练集,另外23252个记录作为测试集。

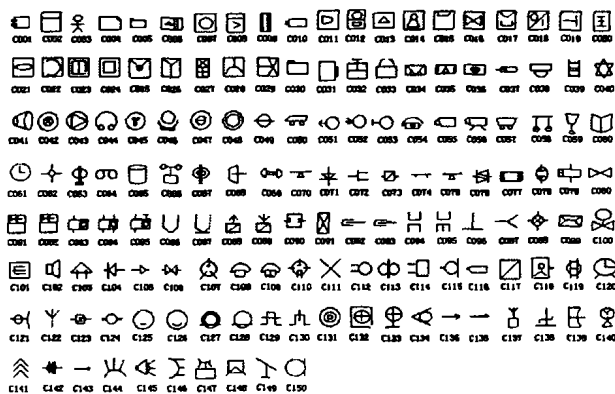
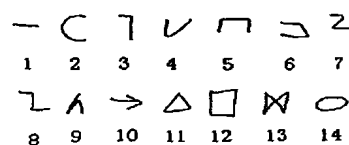


图6 实验图形集

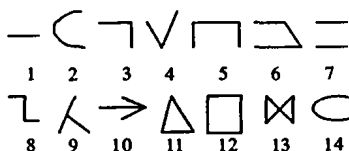
4.2 实验结果分析

我们的实验都是在 Intel P4 PC (1.4G Hz CPU, 768MB 内存), Microsoft Windows XP 环境下进行的。在实验中,我们对 SVM 重复学习、Syed 等^[14]提出的增量学习方法、我们前期所提出的增量学习方法 α -ISVM^[13]和本文提出的增量式主动学习算法进行了笔划识别的训练时间、识别精度和样本数的比较,如图8所示。实验证明:在具有相当精度的情况下, SVM 增量式主动学习算法的训练时间平均仅仅为重复学习方法的1/16、 α -ISVM 算法的1/6、Syed 等的算法的1/4,增量学习过程中实际上增加的样本数目得到大幅度下降,而且可以有效鉴别不同用户的笔划输入。这说明本文所提出的 SVM 增量式主动学习方法能自动挑选重要样本,并在训练少量样

本的情况下就能实现鉴别用户和较好的识别精度,且无需 α -ISVM 算法那样加入误分样本。



(a) 原始笔划类别



(b) 规整化笔划类别

图7 手绘草图的常用笔划集

我们将基于 SVM 的笔划识别结果应用于用户手绘习惯模型的草图图形识别。在实验中对收集到的草图图形数据分别用每种画法的前1、3、5、7个样本(表示为 T_1 、 T_3 、 T_5 和 T_7)进行训练,并取预测列表候选图形的前1、3、5、7个类别为判据确定其正确率,实验结果如表1所示。通过实验可以看出:采用用户手绘习惯模型进行草图识别仅需要很少量的训练数据就能够达到很好的识别效果,而且随着训练样本数的增加,识别精度也有相应的提高。同时,与精确匹配方式相比,模糊匹配具有明显优势,尤其是在训练样本较少的情况下,主要原因在于:训练样本在起始阶段不能穷举所有的情况。

总结与展望 用户适应性是在线手绘草图识别系统中一个十分重要的课题,尤其对于复杂草图的识别来讲更加困难。本文基于前期手绘草图识别系统的研究和开发^[8,9],对与用户适应性相关的两个重要问题进行了研究。一是采用转角函数为笔划特征,使用支撑向量机(SVM)主动式增量学习方法对笔划进行识别,与传统基于支撑向量机的增量学习方法相比,该方法可以充分利用笔划的历史数据,并能有效“分析”用户数据,选择更为重要的数据来训练支撑向量机,既避免了重复训练的开销,更可以在保持精度的前提下,减少人工标注的工作量。二是设计了增量决策树来构造用户手绘习惯模型,该模型充分利用用户输入的笔划构成的几何特征信息、笔划间的空间关系和时序信息,可有效捕捉用户输入习惯和绘图的过程。

程信息;并采用模糊匹配方法对用户输入的复杂草图进行预测,可在用户输入草图信息不完整时为快速草图输入提供帮助。文中所提出的方法取得了很好的实验效果。用户适应性问题是一个涉及人类视觉认知规则、思维习惯和信息模型的复杂课题,本文的研究只是在这一领域的一个简单尝试,我们将

在 SVM 的多值复杂草图分类、基于决策树的学习方法及其在草图识别中的应用、基于隐马尔可夫模型(HMM)的手绘者识别等方面进行更加深入的研究和开发,力求找到解决在线草图识别用户适应性的系统解决方案。

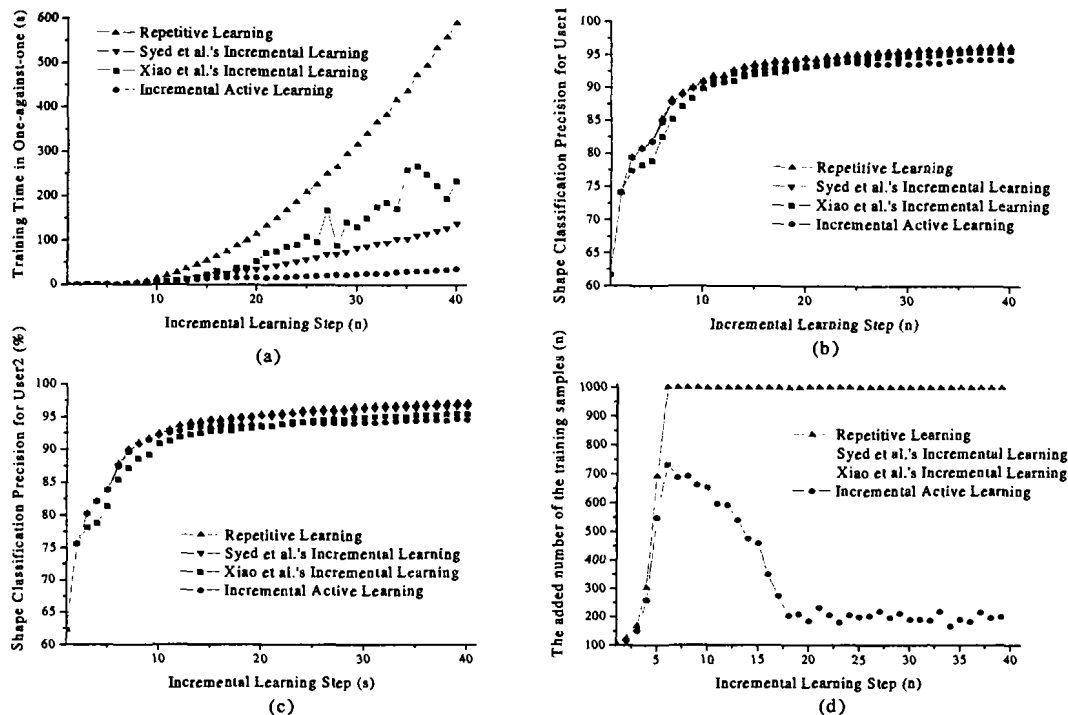


图8 四种 SVM 学习算法识别效果比较

表1 顺序模式下的精确匹配和模糊匹配识别正确率(%)

	预测数据(候选草图图形中的前 n 个)							
	Top 1		Top 3		Top 5		Top 7	
训练数据(前 n 个样本)	精确匹配	模糊匹配	精确匹配	模糊匹配	精确匹配	模糊匹配	精确匹配	模糊匹配
T ₁ (共537个样本)	47.61	77.72	55.49	88.08	58.29	91.37	60.34	92.68
T ₃ (共1611个样本)	76.16	81.81	81.87	92.99	85.16	95.47	86.10	96.96
T ₅ (共2685个样本)	85.35	83.43	89.26	93.67	91.19	96.46	91.50	97.77
T ₇ (共3759个样本)	89.14	85.23	92.43	94.60	93.92	96.96	94.29	98.63

参考文献

- 1 Fish J, Scrivener S. Amplifying the mind's eye; Sketching and visual cognition[J]. Leonardo, 1990, 23(1): 117~126
- 2 Ivan S. Sketchpad: a man-machine graphical communication system[A]. In: Proc. of the 1963 Spring Joint Computer Conf. Baltimore, MD: Spartan Books, 1963. 45~53
- 3 Herot C F. Graphical input through machine recognition of sketches[A]. In: Proc. of SIGGRAPH'76, 1976. 97~102
- 4 Negroponte N. Recent advances in sketch recognition[A]. AFIPS (American Federation of Information Processing) National Computer Conference, Boston, MA, 1973(42): 663~675
- 5 周若鸿, 孙正兴, 张莉莎, 徐晓刚. 草图理解技术研究进展. 计算机科学, 2004, 31(4): 140~146
- 6 Rabiner L R. Tutorial on Hidden Markov Model and Selected Applications in Speech Recognition[A]. Proc. IEEE, 1989, 77(2): 257~285
- 7 Plamondon R, Srihari S N. On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey[J]. IEEE Transaction On Pattern Analysis and Machine Intelligence, 2000, 22(1): 63~84
- 8 Liu W Y, Jin X Y, Sun Z X. Sketch-Based User Interface for In-

- putting Graphic Objects on Small Screen Device[J]. Lecture Notes in Computer Science, Springer, 2002, 2390: 67~85
- 9 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报, 2003, 已录用
- 10 Dean R. Specifying gestures by example[A]. In Computer Graphics (SIGGRAPH), ACM SIGGRAPH, Addison Wesley, July 1991. 329~337
- 11 Peng B B, Sun Z X, Xu X X. An SVM-based incremental learning for user adaptation in the shape recognized system[A]. In: Proc. of the First Intl. Conf. on Machine Learning and Cybernetics, Beijing, 2002. 1379~1386
- 12 Esther M A. An Efficient Computable Metric for Comparing Polygonal Shapes [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1991, 13(3): 209~216
- 13 萧嵘, 王继成, 孙正兴, 张福炎. 一种 SVM 增量学习算法 α -ISVM. 软件学报, 2001, 12(12): 1818~1824
- 14 Syed N, Liu H, Sung K K. Incremental Learning with Support Vector Machines[A]. In: Proc. Workshop on Support Vector Machines at the Intl. Joint Conf. on Artificial Intelligence (IJCAI-99), Stockholm, Sweden, 1999
- 15 彭彬彬. 草图理解中的用户适应性研究: [南京大学申请硕士学位论文]. 50~51