

基于模板匹配和 SVM 的草图符号自适应识别方法

姜映映^{1),2)} 田 丰¹⁾ 王绪刚¹⁾ 戴国忠^{1),2)}

¹⁾(中国科学院软件研究所人机交互技术与智能信息处理实验室 北京 100190)

²⁾(中国科学院软件研究所计算机科学国家重点实验室 北京 100190)

摘 要 在草图符号的自适应学习中,不同用户的训练样本数量可能不同,保持在不同样本数量下良好的学习效果成为需要解决的一个重要问题.提出一种自适应的草图符号识别方法,该方法采用与训练样本个数相关的分类器组合策略将模板匹配方法和 SVM 统计分类方法进行了高效组合.它通过利用支持小样本学习的模板匹配方法和支持大量样本学习的 SVM 方法,并同时利用草图符号中的在线信息和离线信息,实现了不同样本个数下自适应的符号学习和识别.基于该方法,文中设计并实现了支持自适应识别的草图符号组件.最后,利用扩展的 PIBG Toolkit 开发出原型系统 IdeaNote.评估表明,该方法可以在 24 类草图符号分别使用 1 到 20 个训练样本时具有较高的识别正确率和较好的时间性能.

关键词 符号识别;自适应学习;模板匹配;SVM;分类器组合;组件

中图法分类号 TP391

DOI 号: 10.3724/SP.J.1016.2009.00252

Adaptive Symbol Recognition for Sketch-Based Interfaces Based on Template Matching and SVM

JIANG Ying-Ying^{1),2)} TIAN Feng¹⁾ WANG Xu-Gang¹⁾ DAI Guo-Zhong^{1),2)}

¹⁾(Intelligence Engineering Laboratory, Institute of Software, Chinese Academy of Sciences, Beijing 100190)

²⁾(State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190)

Abstract During adaptive learning of symbols in sketch-based interfaces, the number of training samples may be different for different users and it is challenging for recognition methods to learn with flexible sample numbers. This paper proposes an adaptive symbol recognition method for sketch-based interfaces. It combines template matching method that could learn with few samples and SVM method that could learn with more samples by a strategy related to sample numbers. Both online information and offline information are utilized. Thus it could learn and recognize with different sample numbers. Based on the proposed method, the authors build a symbol widget that supports adaptive recognition. At last, a prototype system, IdeaNote, is built based on the extended PIBG Toolkit. Evaluation shows that when there are 24 kinds of symbols, the method could achieve high recognition accuracy and good time performance with 1 to 20 training samples.

Keywords symbol recognition; adaptive learning; template matching; SVM; classifier combination; widget

收稿日期:2008-07-29;最终修改稿收到日期:2008-09-10. 本课题得到国家自然科学基金(U0735004,60503054,60603073)、国家“八六三”高技术研究发展计划项目基金(2007AA01Z158)资助. 姜映映,女,1982年生,博士研究生,主要研究方向为人机交互、笔迹计算、多通道交互. E-mail: jyy@iel.iscas.ac.cn. 田 丰,男,1976年生,博士,副研究员,主要研究方向为笔式交互、虚拟现实和用户界面工具. 王绪刚,男,1976年生,博士,助理研究员,主要研究方向为人机交互和机器学习. 戴国忠,男,1944年生,研究员,博士生导师,主要研究领域为人机交互、软件工程和计算机图形学.

1 引言

手绘草图是人类一种自然而直接的思路外化和交流方式^[1], 它可以拉近思维空间与物理空间的距离并辅助短时记忆, 帮助我们解决问题. 随着笔式计算设备的普及, 出现了一些著名的草图系统, 如 Silk^[2] 等, 它们需要草图识别和理解技术的支持. 然而, 已有的草图系统大都预先定义了可以识别的草图符号的范围和符号的形状, 用户只能按照系统规定的方式绘制系统所包含的符号. 这使得用户在使用时往往要放弃自己的一些个性化习惯, 违背了自然人机交互技术的原则. 支持自适应学习的草图识别和理解技术可以一定程度上解决这个问题. 它试图从用户的样本数据中学习用户的习惯, 提高系统对特定用户的识别率. 由此带来的一个重要问题是不同用户添加的训练样本数量可能不同, 草图符号自适应识别方法需要在训练样本数量多或者少的时候都能很好地学习草图符号.

本文提出一种基于模板匹配和 SVM 的草图符号自适应识别方法, 它通过组合模板匹配方法和 SVM 统计分类方法, 可以有效支持不同训练样本个数下的草图符号学习. 为提高识别正确率, 该方法利用了草图符号中的在线笔划信息与离线图像信息, 将在线模板匹配和离线模板匹配组合并将在在线 SVM 分类器和离线 SVM 分类器组合. 基于自适应草图符号识别方法, 本文设计并实现了支持自适应识别的草图符号组件, 并利用扩展的 PIBG Toolkit^[3] 构造了原型系统 IdeaNote. 评估表明, 当训练样本个数多变时, 该方法具有良好的自适应特性, 可以针对不同用户输入的样本个数动态调整学习和识别算法, 从而可以始终保证较高的识别率和好的时间性能.

本文第 2 节介绍相关研究; 第 3 节给出该自适应识别方法的详细介绍, 包括自适应识别方法的框架和该方法的关键技术: 预处理、分类方法以及分类器组合; 第 4 节介绍基于自适应识别方法的组件和应用; 第 5 节对该方法进行评估; 第 6 节总结全文.

2 相关研究

草图符号识别是草图识别和理解的基础. 由于草图符号可以具有任意的形状、任意的绘制顺序, 草图符号的识别面临着不同绘制习惯、草图是二维和非时序的等项挑战. 目前, SVM、HMM、神经网络和

贝叶斯网络(BN)等基于统计的方法被广泛应用于草图符号识别领域. SVM 具有比神经网络、HMM 更高的识别率^[4]. 具体到自适应符号识别, Sun^[5] 研究了基于 HMM 的自适应在线草图识别方法, 通过动态改变 HMM 的状态个数识别用户的绘制习惯. Zhang^[6] 则比较了 SVM、HMM 和 BN 用于自适应草图识别的准确率和训练时间. 然而, 上述方法并不适合用少量几个样本定义新的草图符号类别.

目前已有一些研究尝试解决小量训练样本情况下的草图符号学习问题. Veselova^[7] 试图通过单个样本给出符号的定义, 样本的泛化能力来源于对人的感知特性的利用. 该方法目前只支持由线和椭圆组成的符号, 且符号的识别与线和椭圆的个数有关. Hammond^[8] 定义了一种 Ladder 描述语言, 它可以定义草图如何被绘制、显示和编辑. 但通过语言描述定义草图符号的方法并不直观, 即使简单的草图符号也需要较长的描述, 并不适合草图符号的用户自定义. Kara^[9] 提出一种基于图像的模板匹配方法识别手绘草图符号. 它支持单样本学习, 但训练样本(模板)个数增多时, 算法的识别时间性能会下降. Wobbrock^[10] 提出一个易于实现的 \$1 手势识别器. 它可以在少量样本下取得较高的识别正确率, 但它只支持单笔划手势的识别.

草图符号的识别方法主要有两类: 离线识别方法和在线识别方法. 离线识别适于处理具有不规范书写顺序、有重叠笔划的符号, 但不适合书写不正规的符号; 在线识别适于处理书写扭曲(distorted)、正常书写顺序的符号, 但不适合处理不规范书写顺序、重叠笔划(overwriting)的符号^[11]. 在手写字符识别中, 离线信息和在线信息被同时利用以提高识别方法的识别率^[11-12]. Keshari^[12] 采用分类器组合的方式, 分别对在线信息和离线信息用 SVM 建模.

本文将利用离线信息和在线信息, 组合模板匹配方法和 SVM 方法来识别草图符号. 离线信息与在线信息的同时利用可以提高识别方法的识别正确率, 模板匹配方法与 SVM 方法的组合适于处理不同训练样本个数下的符号学习, 并具有较好的时间性能.

3 基于模板匹配和 SVM 的自适应识别方法

3.1 基于模板匹配和 SVM 的自适应识别方法框架

图 1 给出了基于模板匹配和 SVM 的草图符号自适应识别方法的框架. 笔迹输入首先被预处理, 包括在线预处理和离线预处理. 若处于训练状态, 则训

练在线和离线 SVM 分类器,并添加模板到模板库.模板包括类别标识、对应的图像和笔划.若处于识别状态,则分别用在线和离线 SVM 进行识别,并把待分类样本的笔划和图像与模板库中的模板进行在线和离线匹配.在线 SVM 分类器、离线 SVM 分类器

以及两种 SVM 的组合构成混合 SVM 分类方法;在线模板匹配、离线模板匹配以及两种模板匹配的组合构成混合模板匹配方法.最后根据训练样本的个数,对混合 SVM 方法和混合模板匹配方法进行组合得到最终的分类结果.

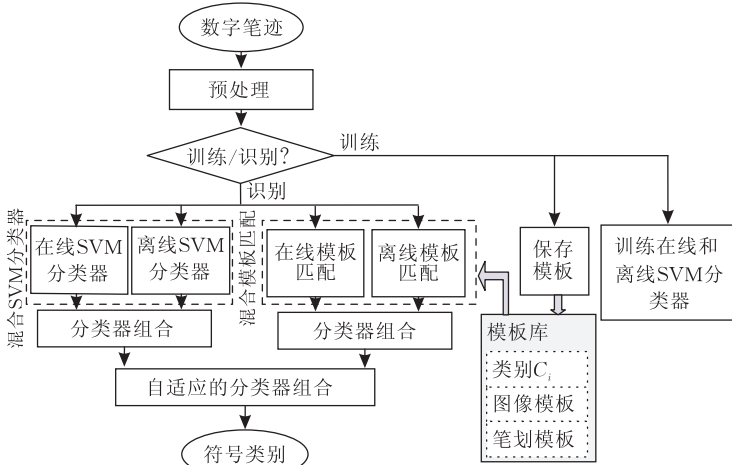


图 1 基于模板匹配和 SVM 的草图符号自适应识别方法框架

图 1 中的草图符号自适应识别方法的关键技术包括预处理(含在线和离线预处理)、支持不同样本个数学习的分类方法(含在线和离线模板匹配以及在线和离线 SVM 分类器)以及基于加权求和的分类器组合技术.下面将分别介绍这 3 个方面.

3.2 预处理

由于手绘草图符号具有很大的随意性,草图符号可能具有不同的大小和不同的笔划数.预处理将草图符号通过放缩、旋转、平移等变换得到归一化后的草图符号,以消除同类别不同符号之间的差异.由于本文同时用到了在线信息和离线信息,预处理包括相应的两部分:在线预处理和离线预处理.

在线预处理首先将多笔划表示的草图符号转为单笔划表示,即将原始符号的多个笔划连接成单笔划(第 i 个笔划的末点与第 $i+1$ 个笔划的起点连接).然后将符号平移、放缩并重采样.预处理后的草图符号用单笔划表示,具有相同的包围盒大小以及相同的采样点个数.

离线预处理首先将原始符号平移、放缩.然后,

对平移、放缩后的符号中的每个笔划重采样,使得重采样后笔划内相邻点之间的距离为 1. 最后,将用笔划表示的草图符号转存为图像表示.同时,采用距离变换^[13]的方法计算图像的距离矩阵.图像的距离矩阵储存着图像中任一像素点距离图像中黑色像素点的最近距离. Kara 提出的基于图像模板匹配的草图符号识别方法^[9]采用了与本文类似的离线预处理方法.

图 2 给出草图符号预处理结果的一个实例.图 2 (a)为原始的草图符号,包含两个笔划(笔划起始点用大点表示),共有 60 个采样点.原始符号包围盒左上角坐标为(64,90),右下角坐标为(150,184).图 2 (b)~(d)分别为由图 2(a)预处理得到的模板.图 2 (b)为 48×48 大小的图像模板.图 2(c)为用灰度图像表示的距离矩阵模板,大小为 48×48.灰度值描述了距离的远近,颜色越深的点与对应图像中黑色像素点的距离越远.图 2(d)为在线预处理得到的笔划模板,笔划模板包含 64 个采样点,包围盒大小为 20×20.

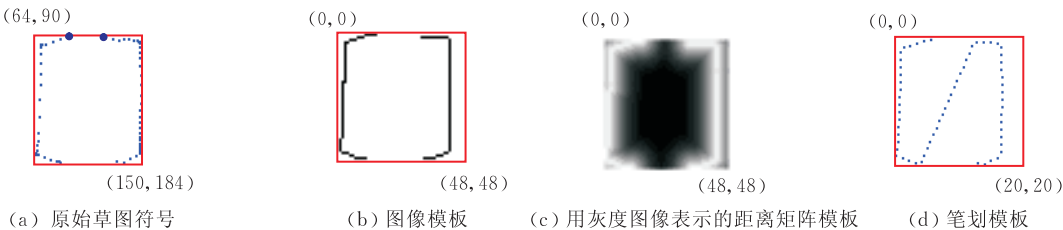


图 2 预处理结果实例

3.3 面向不同样本个数学习的分类方法

3.3.1 支持小样本学习的模板匹配方法设计

模板匹配方法适合小样本学习, 它将当前待分类样本与模板库中的模板进行匹配, 得到待分类样本与模板之间的距离. 为了满足分类器组合的需要, 需将匹配距离转换为识别可信度.

(1) 在线模板匹配距离和离线模板匹配距离

由于笔划对应点间的欧式距离以及图像间的 Hausdorff 距离用于模板匹配具有较高的识别率^[9-10], 本文采用这两种距离用于在线模板匹配和离线模板匹配. 式(1)给出两个笔划 A 、 B 对应点之间的平均欧式距离的定义.

$$d_s(A, B) = \frac{\sum_{i=1}^N \sqrt{(x_{Ai} - x_{Bi})^2 + (y_{Ai} - y_{Bi})^2}}{N} \quad (1)$$

其中, A 、 B 均具有 N 个采样点, (x_{Ai}, y_{Ai}) 和 (x_{Bi}, y_{Bi}) 分别为 A 和 B 中的第 i 个点的坐标.

两个点集合 A 、 B 的 Hausdorff 距离^[14] 定义如下:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (2)$$

其中, $h(A, B) = \max_{a \in A} (\min_{b \in B} \|a - b\|)$, $h(B, A) = \max_{b \in B} (\min_{a \in A} \|a - b\|)$, a 为集合 A 中的点, b 为集合 B 中的点.

由于预处理得到的图像的距离矩阵储存着图像中任一像素点距离图像中黑色像素点的最近距离, 计算一幅图像上某一点到另一幅图像上黑色像素点的最小距离可以通过叠加一幅图像到另一幅图像的距离矩阵上直接得到, 从而加快 Hausdorff 距离的计算速度.

(2) 匹配距离到识别可信度的转换

为方便模板匹配方法与 SVM 分类器的组合, 将模板匹配得到的距离转化为识别可信度. 假设 $d_1(A, B) = d_s(A, B)$, $d_2(A, B) = H(A, B)$. 首先将匹配距离归一化, 即

$$d'_i(A, B) = \frac{d_i(A, B) - dMin_i}{dMax_i - dMin_i}, \quad i=1, 2 \quad (3)$$

其中, $dMax_i$ 为第 i 种模板匹配方法对应的最大匹配距离, $dMin_i$ 为第 i 种模板匹配方法对应的最小匹配距离.

采用第 i 种模板匹配方法, 待分类样本 S 与类别为 c_j 的草图符号对应的模板的最近距离为

$$d_i(c_j | S) = \min\{d'_i(T, S_i) | T \in \text{templateLib}_i, T.classType = c_j\}, \quad i=1, 2; j=1, 2, \dots, CN \quad (4)$$

其中, S_i 为由待分类样本 S 预处理得到的对应于第 i

类模板匹配方法的模板, templateLib_i 为对应于第 i 类模板匹配方法的模板库, $T.classType$ 为模板 T 的类别, cN 为草图符号的类别总数.

式(5)给出第 i 种模板匹配方法下, 待分类样本 S 的类别为 c_j 的识别可信度的计算公式. 具有最高识别可信度的类别为最终的分类结果.

$$p_{Ti}(c_j | S) = 1 - d_i(c_j | S), \quad i=1, 2; j=1, 2, \dots, cN \quad (5)$$

3.3.2 支持大样本学习的 SVM 分类器设计

SVM 分类器适合训练样本较多情况下的符号学习. 本文采用支持多类分类的 SVM 分类器, SVM 的输出为待分类样本被识别为每个类别的识别可信度. 具有更高识别可信度的符号类别更有可能是草图符号的最终类别. SVM 分类器由在线 SVM 分类器和离线 SVM 分类器组成.

在线 SVM 分类器的输入为在线的笔划信息. 假设预处理得到的表示草图符号的笔划为 $S' = \{(x_1, y_1), (x_2, y_2), \dots, (x_{PN}, y_{PN})\}$, (x_i, y_i) 为笔划 S' 的第 i 个点, PN 为笔划包含的点的个数. 则用于在线分类的特征向量为

$$\mathbf{F}_{\text{online}} = \{\sin_1, \cos_1, \sin_2, \cos_2, \dots, \sin_{PN-1}, \cos_{PN-1}\} \quad (6)$$

其中,

$$\sin_i = \frac{y_{i+1} - y_i}{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}},$$

$$\cos_i = \frac{x_{i+1} - x_i}{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}.$$

离线 SVM 分类器的输入特征向量从图像信息中得到. 将预处理得到的图像模板分为 $m \times m$ 个小块, 计算每一小块图像中黑色像素点占黑白像素总数的比例, 得到 $m \times m$ 维的输入向量.

$$\mathbf{F}_{\text{offline}} = \{\text{density}_1, \text{density}_2, \dots, \text{density}_{m \times m}\} \quad (7)$$

其中, density_i 为第 i 个图像块中黑色像素数占黑白像素总数的比例.

在线 SVM 分类器的输出为 $p_{S1}(c_j | S)$, $j=1, 2, \dots, CN$. CN 为草图符号的类别总数. 离线 SVM 分类器的输出为 $p_{S2}(c_j | S)$, $j=1, 2, \dots, CN$.

3.4 基于加权求和的分类器组合

分类器组合可以结合不同分类器的优点, 从而提高分类效果. 基于求和的分类器组合策略优于其它的分类器组合策略^[15], 本文采用基于加权求和的分类器组合方式.

3.4.1 分类器组合系数的计算

在线识别和离线识别组合时的组合系数设置在 0.5 和 0.6 之间最合适^[8], 我们设置在线模板匹配和离线模板匹配的分类器组合系数为 $a_{T1} = a_{T2} = 0.5$, 它们构成了混合模板匹配方法; 设置在线 SVM 分类器和离线 SVM 分类器的分类器组合系数为 $a_{S1} = a_{S2} = 0.5$, 它们构成了混合 SVM 方法。

由于模板匹配方法和 SVM 分类器在不同训

练样本个数下的表现不同, 不适合采用组合权重在分类器间平均分配的方式. 因此, 采用训练样本个数相关的分类器组合策略对它们进行组合. 图 3 给出了混合模板匹配和混合 SVM 的组合策略. 其中, $thres1$ 和 $thres2$ 是用于计算组合系数的阈值, $sampleNum$ 为用户的训练样本的个数, $templateNum$ 为用于模板匹配的模板的个数, $svmTrainingNum$ 为 SVM 训练所采用的训练样本的个数。

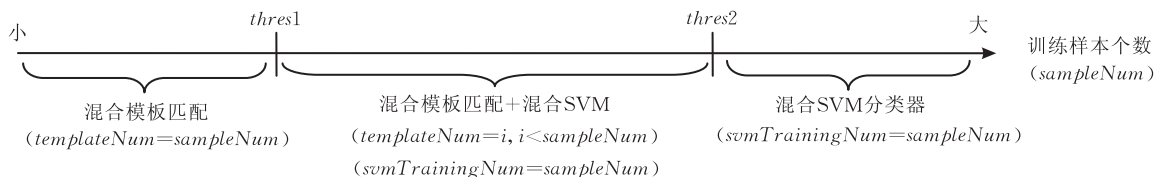


图 3 与训练样本个数相关的分类器组合策略

(1) 当训练样本数小于等于 $thres1$ 时, 采用混合模板匹配的方法进行识别, 用于模板匹配的模板的个数为训练样本的个数;

(2) 当训练样本个数介于 $thres1$ 和 $thres2$ 之间时, 采用混合 SVM 与混合模板匹配组合的方式, 用于模板匹配的模板为最新采集的 $i (i < sampleNum)$ 个训练样本, 采用 i 个训练样本 (模板) 的模板匹配方法进行识别所需的时间小于采用所有训练样本 (模板) 所需的时间, 用于 SVM 训练的样本为用户的所有训练样本;

(3) 当训练样本个数大于 $thres2$ 时, 采用混合 SVM 分类器进行分类, 用于 SVM 训练的样本的个数为用户训练样本的个数。

混合模板匹配和混合 SVM 的组合系数 a_1 和 a_2 的定义如下:

$$a_1 = \begin{cases} 1.0, & sampleNum \leq thres1 \\ 1.0 - \frac{sampleNum - thres1}{thres2 - thres1}, & thres1 < sampleNum \leq thres2 \\ 0.0, & sampleNum > thres2 \end{cases}, \quad a_2 = 1 - a_1 \quad (8)$$

3.4.2 分类器组合

分类器组合是指对各分类器的识别可信度加权求和得到组合后的识别可信度。

混合模板匹配方法将在线模板匹配和离线模板匹配组合, 它将待分类样本 S 识别为属于类别 c_j 的可信度为 $p_T(c_j | S)$, 计算公式如下:

$$p_T(c_j | S) = \sum_{i=1}^2 a_{Ti} \times p_{Ti}(c_j | S), \quad j=1, 2, \dots, CN \quad (9)$$

混合 SVM 方法将在线 SVM 和离线 SVM 组合, 它将待分类样本 S 识别为属于类别 c_j 的可信度 $p_S(c_j | S)$, 计算公式如下:

$$p_S(c_j | S) = \sum_{i=1}^2 a_{Si} \times p_{Si}(c_j | S), \quad j=1, 2, \dots, CN \quad (10)$$

模板匹配与 SVM 组合的方法将混合模板匹配和混合 SVM 方法进行组合, 得到最终的识别可信度 $p(c_j | S)$ 。

$$p(c_j | S) = \frac{a_1 \times p_T(c_j | S) + a_2 \times p_S(c_j | S)}{Z'}, \quad j=1, 2, \dots, CN \quad (11)$$

Z' 保证所有的 $p(c_j | S)$ 的和为 1, 定义如下:

$$Z' = \sum_{j=1}^{CN} (a_1 \times p_T(c_j | S) + a_2 \times p_S(c_j | S)) \quad (12)$$

具有最高可信度的类别 $classType$ 为待分类样本的符号类别。

$$classType = \{c_i | p(c_i | S) > p(c_j | S), i \neq j, i \in \{1, 2, \dots, CN\}, j \in \{1, 2, \dots, CN\}\} \quad (13)$$

4 基于自适应识别方法的组件和应用

本文基于草图符号自适应识别方法设计实现了一个支持自适应识别的草图符号组件 CSymbol-Widget, 并添加到 PIBG Toolkit^[3] 中. 目前利用扩展的 PIBG Toolkit 已经开发出支持草图符号自定义的原型系统 IdeaNote.

4.1 基于自适应识别方法的组件设计

PIBG Toolkit 是一个基于 PIBG 交互范式思想

的笔式应用的开发平台,开发者可以利用 PIBG Toolkit 来轻松地构造笔式应用. 在 PIBG Toolkit 中主要包括以下几大部分的内容:总体软件体系结构、数据结构定义和管理模块、存储结构定义和管理模块、交互信息处理模块和交互组件库. 本文扩展了 PIBG Toolkit 的交互组件库,即基于自适应草图符号识别方法构造了一个草图符号组件 CSymbolWidget,它支持草图符号的输入和自适应识别.

图 4 给出了草图符号组件 CSymbolWidget 的结构. PIBG Toolkit 的组件主要包括原语产生装置和原语解释装置^[3]. 原语产生装置负责分析接收到的消息,必要时产生交互原语,同时负责进行实时的

词法反馈. 交互原语产生后发送给本组件的原语解释装置. 原语解释装置负责分析得到的交互原语,执行语法层和语义层的处理. 原语解释装置中的原语控制器首先接收到原语产生装置所产生的交互原语. 接着原语控制器来分析该原语是否为当前交互组件接收,若不是,则发送交互原语到其它交互组件. 同时,原语控制器也可以接收由其它组件发来的交互原语. 接着,原语控制器将交互原语发给任务整合器进行原语的整合工作. 任务整合器需要参照当前的上下文信息来辅助任务的构造过程. 同时,在原语解释的过程中,用户可以根据实际的情况来调整原语的解析的结果.

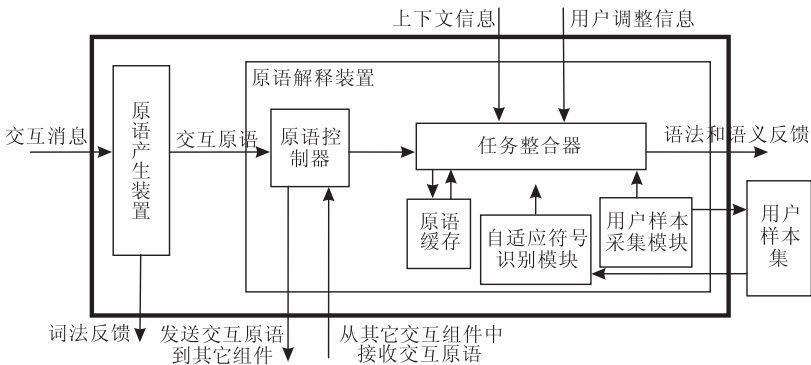


图 4 CSymbolWidget 组件的结构

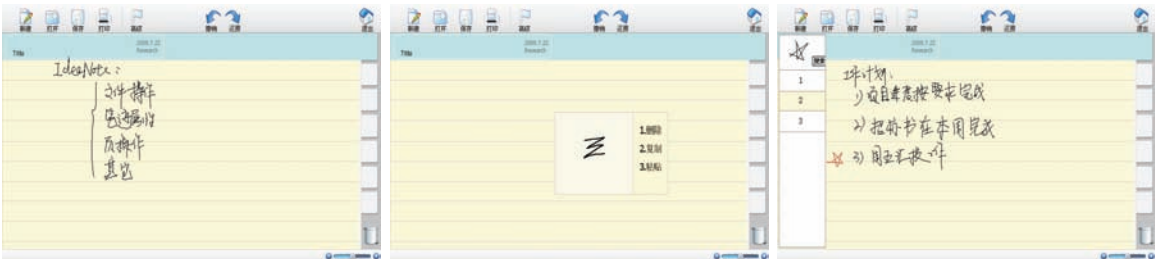
为支持自适应的草图符号识别,原语解释装置包含用户样本采集模块和自适应符号识别模块. 用户样本采集模块用于采集用户的草图符号样本,并存放在用户样本集中. 自适应符号识别模块根据样本采集模块得到的用户样本集,学习用于识别草图符号的分类器. 自适应符号识别模块的核心为上节所述的基于模板匹配和 SVM 的草图符号自适应识别方法,满足在用户样本量大或者小的时候,均能较好地识别用户的草图符号,并满足实时性的要求.

支持自适应识别的草图符号组件 CSymbolWidget 可以针对不同用户的样本个数,快速地构造出相应的草图符号识别算法. 它的引入增加了

PIBG Toolkit 对自适应笔式系统的支持.

4.2 原型系统 IdeaNote

利用扩展的支持自适应草图符号识别的 PIBG Toolkit,可以方便地构造各种支持自适应识别的笔交互系统. 本文实现了支持笔手势和草图符号自定义的原型系统 IdeaNote. IdeaNote 支持用户用数字笔自由地记录笔记和思维过程,笔记编辑主要采用笔手势的交互方式. 由于草图符号自适应识别方法支持单样本学习,系统可以支持笔记搜索等高级功能. IdeaNote 对笔手势和草图符号自定义的支持可以提高系统的可用性,更好地满足用户的个性化需求. 图 5 给出 IdeaNote 的用户界面.



(a) 整体界面 (b) 笔手势自定义界面 (c) 草图符号查询界面

图 5 IdeaNote 用户界面

5 评 估

算法实现时,采用 libsvm^① 中的 SVM 多类分类方法,选用 RBF 核函数,以识别可信度作为输出.式(8)中的参数设置为 $thres1=5, thres2=15$; 训练样本个数在 $thres1$ 和 $thres2$ 之间时,混合模板匹配使用 2 个模板($i=2$).

实验机器的配置为 P4 2.0 CPU, 1GB 内存, wacom 手写屏. 文献[9]中基于特征的识别方法的 24 类测试数据类别被作为评估数据类别(图 6). 3 个用户参与了样本采集,每个用户要完成 480 个草图符号样本(每类符号 20 个样本). 评估结果中的识

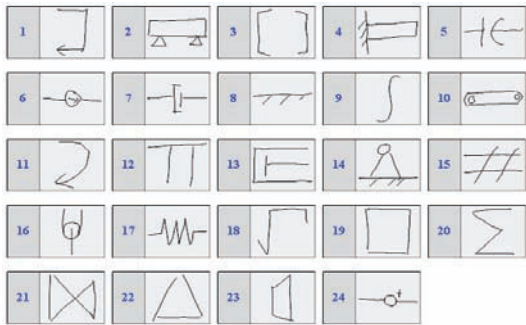
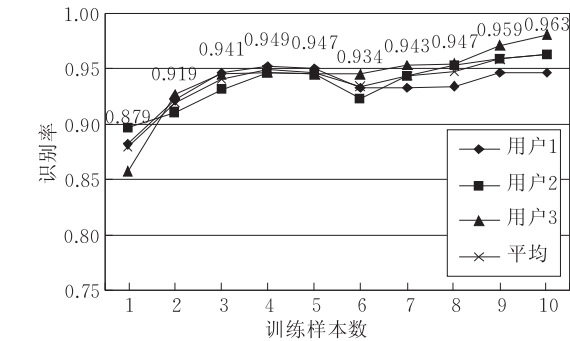
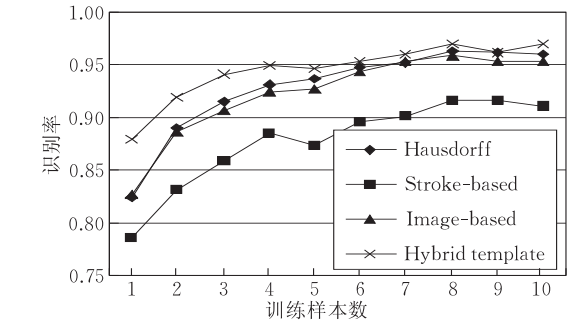


图 6 24 类样本实例



(a) 不同用户的识别率



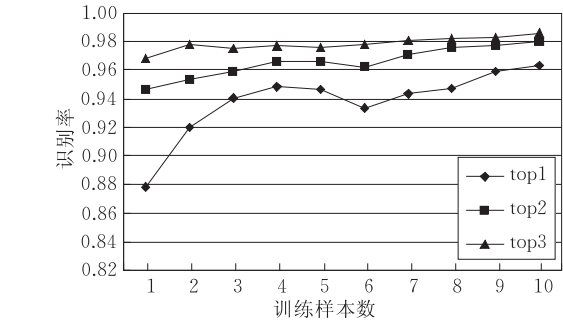
(c) 不同模板匹配方法的识别率

别正确率采用交叉验证的方式计算得到,即从数据集中分别选择不同的样本作为训练样本,剩余的样本作为测试样本得到的识别正确率的平均值.

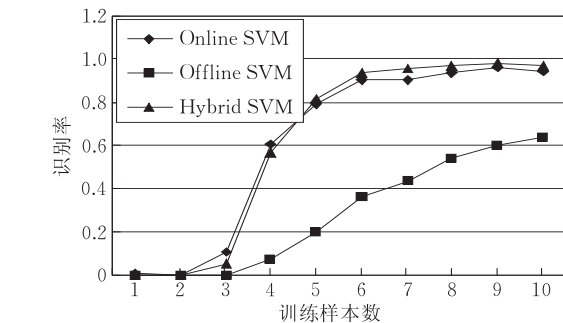
5.1 用户相关的评估

用户相关的评估用每个用户自己的数据进行训练和测试. 图 7 给出用户相关情况下的识别率. 图 7(a)给出本文方法在不同用户下的识别率和 3 个用户的平均识别率. 本文方法在 1 到 10 个训练样本数下的平均识别率为 93.81%,与文献[9]中给出的基于特征的识别方法的识别率 77.78%相比,有明显的提高. 图 7(a)中训练样本个数为 6 时的识别率略低于训练样本个数为 5 时的识别率,原因在于模板匹配(使用 2 个模板)和 SVM 组合的方法在训练样本个数为 6($> thres1$)时的识别率略低于混合模板匹配方法在训练样本个数为 5($= thres1$)时的识别率. 图 7(b)给出了不同模板个数下,正确结果在前一(top1)、前二(top2)和前三(top3)识别候选的准确率(3 个用户的平均正确率). 单模板情况下,正确结果在前 2 个识别候选的比例为 94.69%,在前 3 个识别候选的比例为 96.8%.

图 7(c)和图 7(d)分别给出不同模板匹配方法和不同 SVM 方法的识别率. 混合模板匹配方法



(b) top1, top2, top3 正确率



(d) 不同 SVM 分类方法的识别率

图 7 用户相关的识别率

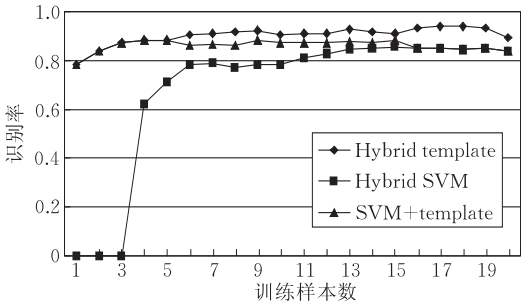
① Chang C C, Lin C J. LIBSVM: A library for support vector machines, 2001. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

(Hybrid template) 的识别率高于离线模板匹配方法 (Hausdorff)、在线模板匹配方法 (stroke-based) 和文献 [9] 中基于图像的模板匹配方法 (image-based); 混合 SVM 方法 (Hybrid SVM) 的识别率高于离线 SVM 分类器 (offline SVM) 和在线 SVM 分类器 (online SVM). 因此, 离线方法与在线方法的组合提高了草图符号的识别率.

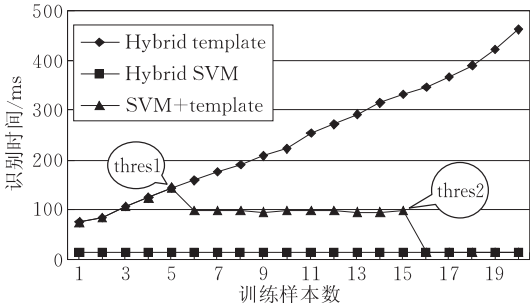
5.2 用户无关的评估

用户无关的评估将所有用户的数据放到一起训练和测试, 图 8 比较了混合模板匹配、混合 SVM 以及模板匹配与 SVM 组合的方法在不同训练样本个数

数下的识别率和识别时间. 图 8(a) 表明, 本文提出的方法在不同训练样本个数下的识别率较为稳定, 可以满足不同训练样本个数下的草图符号学习要求. 图 8(b) 给出 3 种方法下单个样本的识别时间与训练样本个数之间的关系. 可见, 本文方法的识别时间不随训练样本个数的增多而增多. 当训练样本个数小于 $thres1$ 时, 本文方法的识别率和识别时间与混合模板匹配方法相同; 当训练样本个数介于 $thres1$ 和 $thres2$ 之间时, 识别率和识别时间介于混合模板匹配和混合 SVM 之间; 训练样本个数大于 $thres2$ 时, 识别率和识别时间与混合 SVM 相同.



(a) 不同方法的识别率比较



(b) 不同方法下单个样本的识别时间

图 8 用户无关测试下的识别率及识别时间

6 结 论

本文解决了草图符号的自定义问题, 并支持不同用户样本个数下的草图符号自适应学习, 从而可以更好地支持自然的人机交互技术. 自适应的草图符号识别方法采用将支持大量样本学习的分类器和支持小样本学习的分类器组合的方式, 具体采用了模板匹配方法和 SVM 分类器. 并且该方法充分利用了在线信息和离线信息以提高草图符号的识别率. 基于自适应识别方法, 本文构建了一个支持自适应识别的草图符号组件, 添加到 PIBG Toolkit^[3] 中, 扩展了 PIBG Toolkit 对自适应笔式系统的支持, 并基于扩展的 PIBG Toolkit 构造了支持笔手势和草图符号自定义的原型系统 IdeaNote. 评估表明, 该方法可以支持小样本学习和大量样本学习, 并具有较高的识别率和较好的时间性能.

本文提出的方法仍然具有进一步改进的空间, 我们将在以后的工作中加以解决. 首先, 该方法目前假定所有符号类别包含相同的训练样本个数, 而草图符号自适应学习中, 不同符号类别的训练样本个数可能不同. 在未来的工作中, 我们将尝试解决不同类别符号的训练样本个数不同时的符号识别问题.

其次, 改进分类器组合方法是未来工作之一. 目前, 模板匹配和 SVM 组合的阈值 $thres1$ 和 $thres2$ 以及组合时用于匹配的模板个数 i 是根据经验设定的, 将来希望找到一种策略选择最合适的 $thres1$ 、 $thres2$ 和 i 以取得更好的识别率和识别时间. 再次, 将来需要寻找最有代表性的 i 个模板, 以提高模板匹配和 SVM 组合时的识别正确率.

在自适应系统开发方面, 目前已经开发了支持自适应识别的原型系统 IdeaNote. 将来我们将对该原型系统进行评估, 从而为识别方法的改进和实际应用系统的构造提供更多的支持. 基于此, 我们将开发其它支持草图符号自适应学习的草图系统.

参 考 文 献

[1] Fish J, Scrivener S. Amplifying the mind's eye: Sketching and visual cognition. Leonardo, 1990, 23(1): 117-126

[2] Landay J A, Myers B A. Interactive sketching for the early stages of user interface design//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Denver, Colorado, USA, 1995: 43-50

[3] Tian Feng, Qin Yan-Yan et al. Analysis and design on PIBG Toolkit: A pen-based user interface toolkit. Chinese Journal of Computers, 2005, 28(6): 1036-1042 (in Chinese)

- (田丰, 秦严平等. PIBG Toolkit: 一个笔式界面工具箱的分析与设计. 计算机学报, 2005, 28(6): 1036-1042)
- [4] Tapia E, Rojas R. Recognition of on-line handwritten mathematical expressions in the e-chalk system — An extension// Proceedings of the 8th International Conference on Document Analysis and Recognition. Seoul, Korea, 2005: 1206-1210
- [5] Sun Z X, Jiang W, Sun J Y. Adaptive online multi-stroke sketch recognition based on hidden markov model//Proceedings of the 4th International Conference on Machine Learning and Cybernetics. Guangzhou, China, 2005: 948-957
- [6] Zhang L S, Sun Z X. An experimental comparison of machine learning for adaptive sketch recognition. Applied Mathematics and Computation, 2007, 185(2): 1138-1148
- [7] Veselova O. Perceptually based learning of shape descriptions[M. S. dissertation]. Massachusetts Institute of Technology, Cambridge, 2003
- [8] Hammond T, Davis R. LADDER, a sketching language for user interface developers. Computers & Graphics, 2005, 29(4): 518-532
- [9] Kara L B. Automatic parsing and recognition of hand-drawn sketches for pen-based computer interfaces [Ph. D. dissertation]. Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, 2004

- [10] Wobbrock J O, Wilson A D, Li Y. Gestures without libraries, toolkits or training: A \$1 recognizer for user interface prototypes//Proceedings of the ACM Symposium on User Interface Software and Technology. Newport, Rhode Island, USA, 2007: 159-168
- [11] Tanaka H, Nakajima K, Ishigaki K, Akiyama K, Nakagawa M. Hybrid pen-input character recognition system based on integration of online-offline recognition//Proceedings of the 5th International Conference on Document Analysis and Recognition. Bangalore, India, 1999: 209-212
- [12] Keshari B, Watt S M. Hybrid mathematical symbol recognition using support vector machines//Proceedings of the 9th International Conference on Document Analysis and Recognition. Curitiba, State of Parana, Brazil, 2007: 859-863
- [13] Felzenszwalb P F, Huttenlocher D P. Distance Transforms of Sampled Functions. Cornell2004, Computing and Information Science TR2004-1963, 2004
- [14] Rucklidge W J. Efficient visual recognition using the hausdorff distance//Lecture Notes in computer Science 1173. Springer, 1996
- [15] Kittler J, Hatef M, Duin R P W, Matas J. On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(3): 226-239



JIANG Ying-Ying, born in 1982, Ph. D. candidate. Her major research interests include human computer interaction, ink computing and multimodal interaction.

TIAN Feng, born in 1976, Ph. D., associate professor.

His major research interests include pen-based interaction, virtual reality and user interface tool.

WANG Xu-Gang, born in 1976, Ph. D., assistant professor. His major research interests include human computer interaction and machine learning.

DAI Guo-Zhong, born in 1944, professor, Ph. D. supervisor. His major research interests include human computer interaction, software engineering and computer graphics.

Background

This research is supported by the National Natural Science Foundation of China under grant Nos. 60503054, U0735004, 60603073, and the National High Technology Research and Development Program (863 Program) of China under grant No. 2007AA01Z158.

Sketch-based interface is useful for early design, idea capturing and idea sharing. Currently, there're some popular sketch-based systems, such as Silk, Electronic Cocktail Napkin and MathPad. These systems often predefined symbols that can be recognized, and users need to draw the symbol in the way the system defined. However, as different users often have different preferences for symbol definition and drawing, this predefined way is not very good. To allow users to define their own symbols would be a good solution to this problem.

Adaptive symbol recognition is essential for recognizing users' personalized symbols. When defining a new symbol, different users tend to use different numbers of samples.

This requires the symbol recognition method could learn with adaptive sample numbers. This paper proposes an adaptive symbol recognition method that could learn with either small number of samples or large number of samples. Template matching method that could support small sample learning and SVM that could learn with more samples are combined to achieve adaptive symbol recognition. In addition, both online stroke information and offline image information are utilized to improve the recognition accuracy. A widget built on the proposed method is incorporated into PIBG Toolkit which was developed by the authors' lab and could be used for building pen-based applications. Thus, the extended PIBG Toolkit could support fast building of adaptive pen-based systems. Evaluation shows that the method could achieve high recognition accuracy and good time performance. It could effectively build recognition methods according to users' sample number.