

数字墨水技术：数字世界中的神来之笔

关键词：数字墨水

王长虎 芮 勇
微软亚洲研究院

数字墨水技术

笔和墨，是人类知识和文化传承的重要媒介，是我们从孩提时代就熟悉、掌握并受用一生的工具。千百年来，人们用笔墨记录历史，用笔墨学习，也用笔墨进行艺术创作。从五岳独尊的泰山，到名传四海的黄鹤楼，无不承载着用笔和墨书写的历史。

虽然现在计算机已经非常普及，笔和墨时常被束之高阁，但是人们并没有忘记笔墨对文化的承载，巧妙地开发出了数字墨水技术。这种技术作为一种新的人机界面技术，更方便、也更智能地延续着笔墨的功能。

数字世界中的笔、墨、纸

类似于现实世界中的笔、墨和纸，数字墨水技术也包含数字笔、数字墨水以及承载和显示墨水的屏幕（纸）——触摸屏三个部分。其中笔和触摸屏是硬件，

数字墨水是软件，例如线条的存储、识别与显示技术。

随着触摸屏的逐渐普及，可以预见，数字墨水技术有着广泛的应用前景。学生可以在平板电脑上用手指或笔来涂鸦，不必担心墨水洒得到处都是；商务人士可以快速地在手机或平板电脑上进行会议记录或讨论问题；漫画爱好者可以用笔在电脑上精确绘画，随画随擦。

硬件设备的发展

硬件设备的发展是数字墨水技术的基础。自计算机诞生之日起，研究人员便孜孜不倦地探索更加自然的人机交互模式。

早期的探索 在硬件设备发展初期，数字墨水的输入设备与显示设备是分离的。1957年，汤姆·戴蒙德 (Tom Diamond) 发明了 Stylater 电子写字板。“Stylater”这个名字是由“stylus”(尖笔)与“translator”(翻译)组合而成，它允许用户用

手写笔在电子板上按规则书写，可识别数字或者字母。Stylater 是最早的写字板设备，已具备当前流行的手写笔加写字板的输入模式。

20世纪60年代，兰德公司 (RAND) 的 GRAIL 项目进一步完善了数字墨水设备：一只手写笔、一个写字板、一台显示器以及简单的手写识别技术。GRAIL 强调显示器与手写版笔迹的同步展示与识别，已经具备了数字墨水技术的初步功能。然而，由于受到易用性的限制，它只是一个研究项目，并没有真正转化为产品。

手写输入电脑的商用 从70年代开始，手写输入电脑逐渐进入市场，从 Applicon 公司的 CAD 系统，到 Pencept 公司的 Penpad，再到 GO 公司的 Pen-Point 操作系统以及微软公司的笔式计算、苹果公司的 Newton 系统等。然而，由于这些手写输入系统和电脑体型笨重，并且手写的识别率未达到用户期望，因此在市场上均未取得成功。

1989年,杰夫·霍金斯领导制作了GRiDPad,第一次把输入设备与显示设备合二为一,并在1990年销售了1万多台。GRiDPad的主要用户群是商业公司,并未得到普通消费者认可。

1996年,杰夫·霍金斯再次取得重要突破,他领导的Palm公司推出了一个划时代的产品Pilot 1000,这是第一款真正成功的掌上电脑,同时也为后续的几款更加成功的产品,如Palm V打下了坚实的基础。Palm系列产品体积小可以放在衬衫口袋里,操作简单,价格低廉,可实时手写和识别,受到消费者的欢迎。

智能手机与平板电脑 进入21世纪,智能手机与平板电脑逐渐成为手写输入电脑的中坚力量。微软公司在2002年和2003年分别推出了Windows XP Tablet和Windows Mobile 2003操作系统,利用数字笔和输入面板,用户可以直接在屏幕上写字,并且将自己的手写体方式保存或者转换为文本,输入到其他应用程序中。除了手写文字识别,手写公式识别也将成为可能。

尽管微软公司早期推出的智能手机和平板电脑操作系统优势明显,但随着2007年和2010年iPhone和iPad的推出,史蒂夫·乔布斯领导的苹果公司震撼了全球,成为消费者所拥戴的赢家。乔布斯完全抛弃了物理键盘,甚至抛弃了数字笔,用户可以直接用手指与计算机进行交互,极大地简化了

操作过程。至此,数字笔也不再是数字墨水技术的必需品。

触摸屏外的尝试 乔布斯提出了数字笔的替代品——手指,并使触摸屏和多点触控技术成为智能手机和平板电脑的必备要素。然而,作为数字世界中的“纸”,触摸屏有没有更自然的替代品呢?

2005年,微软亚洲研究院开发了一个名为“万能数字笔”的数字笔原型,集成了微型摄像头、压力传感器、蓝牙模块和存储器芯片,能够把在纸质文档上勾画的线条和文字通过蓝牙同步记录和显示在电脑屏幕上。因此,用户在纸上的书写自动地转为数字化形式,以供进一步识别。唯一的要求是,纸的表面要打印一些淡灰色条纹用于定位万能数字笔的坐标。

近几年,易方公司推出了易方数码笔,利用超音波和红外线对笔尖进行定位,实时地将用户用普通笔芯在普通纸张上书写的线条数字化。除了笔本身,用户只需要将接收单元夹在纸张上部即可。这使得在数字世界的书写和绘画变得与现实世界一样容易。

数字墨水中的“智慧”

硬件方面的进展使得人们在很大程度上可以自由地与机器进行交互,延续着现实世界中笔墨纸的功用。然而,数字墨水技术并不局限于简单的模仿,我们期

望它能在理解用户的所写、所画、所思方面拥有“智慧”,进而提高用户的书写效率,并用数字化技术去整理和归档用户所写,以方便查询和其他应用。这种“智慧”也是现实世界中的笔墨纸所不具备的。

数字和文字识别 早期数字墨水技术关注的是数字和文字的识别,这也是人机交互中最重要、应用最广泛的部分。尽管线下的数字/文字识别技术可以追溯到19世纪,但是由于对硬件的依赖性,早期的在线识别技术往往与硬件的推出相关,并且其识别率在某种程度上决定了硬件设备的成败。

最早的电子写字板Stylater是具备数字/文字识别功能的。在Stylater上有3条铜导线,它们被两个黑点分成了7段,如图1所示。围绕这两个黑点,以特

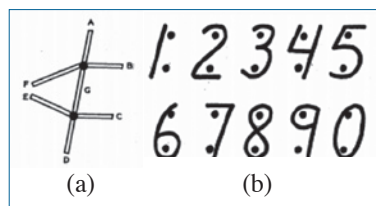


图1 Stylater中的数字识别。(a)被两个黑点分成的7段铜导线,笔尖划过的导线会通电,不同的通电模式对应不同的数字。(b)用两个黑点约束的数字的书写模式。

定模式书写数字,笔尖将划过不同的导线组合,从而记录下不同的通电模式以用于识别数字。同理,用4个黑点便可以识别出不同的字母。

手写识别在早期的手写板

和手写输入电脑中非常重要，GRiDPad 和 Palm 之父杰夫·霍金斯便是以此起家。当他了解到某公司的手写输入识别系统可以要价 100 万美元的时候，他便以极快的速度写了一个自认为更好的识别算法，并因此成为 GRiD-Pad 项目的负责人，而其识别算法成为 GRiDPad 及后续几个项目的核心识别模块。



图2 Graffiti手写字母表

早期的手写识别系统的鲁棒性并不是很高，因此限制了手写输入电脑的实用性。Palm 公司发明了 Graffiti 手写字母表用于手写识别，如图 2 所示。Graffiti 非常简单，每一个字符都有特定的一笔画书写模式，因此下笔和抬笔就意味着一个字符书写完成，非常易于识别。尽管初学者需要学习和适应 Graffiti 的书写模式，但是一旦适应，识别率比之前的无约束系统要高很多。

如今，手写识别技术已经比较成熟。在微软的 Tablet, Windows mobile/phone 以及 Windows 8 系统中，均有鲁棒的手写识别技术，而且不需要类似 Graffiti 的特殊

书写方式，就能够以单词为单位进行识别。

手势识别 伴随着手写数字和文字识别技术的广泛应用和日益成熟，其他相关的手写识别技术也逐渐进入人们的视野。

在早期的硬件中就出现了手势识别技术，例如 20 世纪 60 年代的 GRAIL 项目。手势识别可以看作是在手写板和触摸屏上传递信息和命令的一种方式，现在已经广泛应用于触摸屏设备中。比如，在 Windows Phone 8.1 系统中，手指从屏幕顶端向下滑动，就会出现“通知中心”，来显示设定的主要应用的最新信息。触摸屏上的手势往往非常简单，容易记住，并且在识别后便隐去痕迹。某些手势，比如从右向左画短线，可以用来擦除文字；或者用圆圈聚合某些线条来形成一个形状或单词。

形状和流程图识别 手势

识别实际上也可以看作是简单的形状识别，只是在识别之后便形成特殊的命令。形状识别自动识别用户所画的形状，识别一旦完成便选择用标准的形状来替换。在早期的手写输入系统中，形状识别只是作为流程图的一部分用来编写电脑程序。随着流程图的广泛应用，手写形状识别和复杂流程图识别的研究工作延续至今。

如图 3 所示，在线的复杂流程图识别不仅需要文字和形状识别技术，而且更重要的是，要首先根据线条顺序和空间位置关系，把用户画的线条合理地聚成若干组，使得每一组线条对应一个形状或一组文字。因此，流程图识别以及后续将要介绍的复杂线条识别都需要进行线条的分割，并根据相关领域的先验知识制定策略来简化算法。例如，在文献 [1] 中，流程图的构造规则

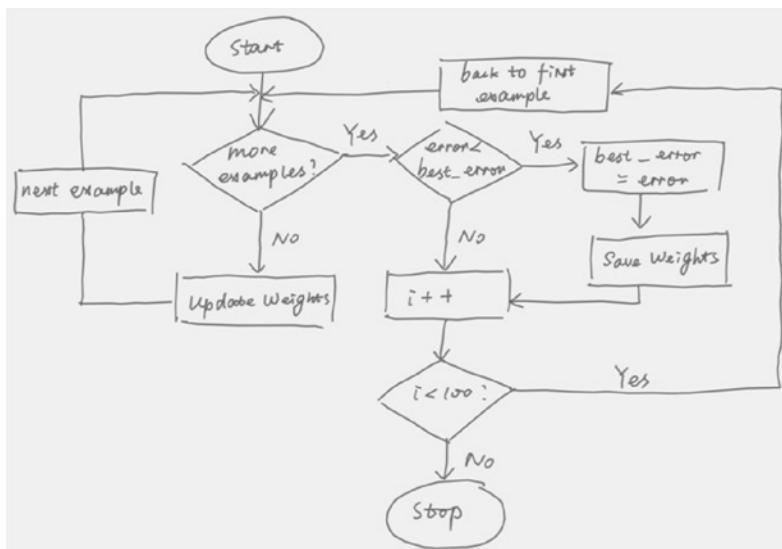


图3 流程图示例^[1]

(例如基本形状之间由箭头连接)对降低算法复杂度起到了至关重要的作用。

其他领域线条图识别 除了流程图识别,在不同领域中均有数字墨水的用武之地。比如,手写公式识别、电路图识别、化学分子图识别、乐谱识别等技术,使得各行各业的人们都可以享受数字墨水的智慧。拿起微软亚洲研究院的万能数字笔,再利用乐谱识别技术,音乐人便可以在作曲的同时欣赏刚刚创作的音乐;化学系的学生画出不同的分子图,并可以实时地在电脑中模拟化学反应;准备毕业论文的研究生也可以直接在 Word 中写出数学公式……

数字墨水足够“聪明”了吗?

我们一方面希望数字墨水技术使我们在数字世界中能够自然地用笔来记录、学习、交流和创作,另一方面期盼能充分利用计算机的智慧,使数字世界中的笔更加智能。数字墨水就像渴望知识的少年,在触摸屏时代,积极地学习知识,茁壮成长。

从写到画 小孩子总喜欢在纸上和墙上涂涂画画,那么,计算机能够识别小孩子画的简单的线条画和彩色画吗?前面介绍的数字墨水识别技术大多依赖很强的领域先验知识,因此算法往往只能应用在相关领域,却无法自动识别孩子们笔下无限可能的物体。



图4 普通物体的线条画示例^[3]

近几年,随着草图搜索技术^[2]的发展,普通物体的草图识别研究也逐渐开展起来。我们通过数据驱动的方式,基于数百万张卡通图片,建立了一个草图识别系统 Sketch2Tag^[3,4],尝试识别任意物体的线条画,如图4所示。希望在不久的将来,数字墨水技术不仅可以鲁棒地识别线条画,还可以识别孩子们的彩色画。

从在线到离线 随着智能手机的普及,用户可以随时用手机拍下自己在纸上记录的内容或者与他在白板上的讨论过程,以备存档和查找。那么,计算机能自动分析用手机拍下的内容并将其数字化吗?



图5 ShadowDraw系统,自动对用户所绘线条进行实时指导。每两行为一个例子,其中第一行是呈现给用户的结果,第二行是算法产生的用于推荐的阴影

在线的手写线条是相对完整的,并具有时间信息,可以用来简化复杂线条的分割与识别。然而,离线拍下的线条没有时间信息,并且从背景中提取出来后可能变得不完整。那么,如何设计有效快速的识别算法来识别离线线条,成为数字墨水技术需要攻克难题。

更智能 当前数字墨水识别技术大多可以识别用户手写的线条,并可以用标准的文字或形状替代。然而,如果遇到极其相似的形状还能有效区分吗?允许用户随意修改吗?它会猜测用户的想法,并补全用户未画的部分吗?

研究人员一直期望并努力使数字墨水技术变得更加智能。我们建立了一个名为 SmartVisio^[5]的系统,能有效区分极其相似或相关的形状,例如圆角矩形和尖角矩形,并允许用户自由修改编辑。为了让计算机帮助用户绘画,微软雷德蒙研究院建立了一个名

CCF YOCSEF走进金山与智谷

CCF YOCSEF 6月的 CLUB 活动是走进金山与智谷。YOCSEF 荣誉委员侯紫峰、李方、刘挺，主席胡春明，副主席韩银和、祝烈煌，学术秘书谭晓生、田丰，以及其他委员等 40 多人参加此次活动。

金山公司 CEO 张宏江博士及智谷公司创始人、总裁林鹏先生代表智谷公司向与会嘉宾表示欢迎。金山集团副总裁刘伟先生向大家介绍了拥有 25 年历史的金山公司及其最新发展。由雷军先生担任董事长、张宏江博士担任 CEO 的金山公司，日前已成功地从传统的软件公司转型为互联网公司。刘伟用妙趣横生的方式介绍了金山“3+1”业务模式。

林鹏向大家介绍了智谷公司，并与大家分享了国际化专利博弈的那些事儿以及高质量核心专利的真实情况。智谷公司是一家以促进创新和发明来推动中国原创技术发展的高科技公司，将发明视为核心竞争力。智谷致力于在中国引领一种以创新为驱动的经济形态，通过营造一种将发明高效地转化成产品的市场环境，最大化地挖掘原创技术的经济价值。



为 ShadowDraw^[6] 的系统，其在用户画线条的同时，在背景中显示阴影，阴影越深表明越可能出现线条，从而指导用户绘画，如图 5 所示。

数字墨水技术，就像一只神来之笔，为我们在数字世界中延续那笔尖划过纸张的感觉。更加自然，更加智能，将是数字墨水技术不懈追求的目标。■



王长虎

微软亚洲研究院研究员。主要研究方向为新一代多媒体搜索、视觉识别、草图搜索与理解等。chw@microsoft.com



芮勇

CCF 会员。微软亚洲研究院首席研究员。主要研究方向为移动计算、无缝漫游及移动互动等。yongrui@microsoft.com

参考文献

- [1] A. Lemaitre, H. Mouchere, J. Camillerapp, and et al.. Interest of syntactic knowledge for on-line flowchart recognition. In *Graphics Recognition. New Trends and Challenges*. Springer, 2013: 89~98.
- [2] 王长虎, 张磊. 草图搜索的魅力与挑战. 中国计算机学会通讯, 2012;12(8).
- [3] Z. Sun, C. Wang, L. Zhang, and et al.. Query-adaptive shape topic mining for hand-drawn sketch recognition. *Proceedings of the 20th ACM international conference on Multimedia*. ACM, 2012: 519~528.
- [4] Z. Sun, C. Wang, L. Zhang, and et al.. Free hand-drawn sketch segmentation. *ECCV 2012*. Springer, 2012: 626~639.
- [5] J. Wu, C. Wang, L. Zhang, and et al.. Sketch Recognition with Natural Correction and Editing. *AAAI 2014*.
- [6] Yong Jae Lee, Larry Zitnick, and Michael Cohen. *ACM Transactions on Graphics*, 2011.