

基于时序特征的草图识别方法

于美玉 吴昊 郭晓燕 贾棋 郭禾
(大连理工大学软件学院 辽宁 大连 116621)

摘要 草图识别是一项很具有挑战性的工作。目前,大部分草图识别的工作都将草图当作普通的纹理图像,忽视了草图的时序性。因此,文中通过挖掘草图的时序性,将草图笔画按照时间分组。为进一步利用时序特征在草图识别过程中的作用,使用了循环神经网络将笔画分组按照时间序列作为输入,最后使用联合贝叶斯将各个时序下获得的草图特征进行整合,完成草图的识别工作。在公开标准数据集上对所提算法进行了测试,实验结果显示该算法的识别准确率明显高于其他算法。

关键词 草图识别,时序性,循环神经网络,门控制单元,联合贝叶斯

中图分类号 TP391 **文献标识码** A

Sequential Feature Based Sketch Recognition

YU Mei-yu WU Hao GUO Xiao-yan JIA Qi GUO He

(School of Software Technology, Dalian University of Technology, Dalian, Liaoning 116621, China)

Abstract Recognizing freehand sketches is a greatly challenging work. Most existing methods treat sketches as traditional texture images with fixed structural ordering and ignore the temporality of sketch. In this paper, a novel sketch recognition method was proposed based on the sequence of sketch. Strokes are divided into groups and their features are fed into recurrent neural network to make use of the temporality. The features from each temporality are combined to produce the final classification results. The proposed algorithm was tested on a benchmark, and the recognition rate is far above other methods.

Keywords Sketch recognition, Temporality, Recurrent neural network, Gate recurrent units (GRU), Joint bayes

1 引言

自古以来,草图就是人类记录和沟通的重要手段。在现代应用中,用户可以通过画出不认识或者想买的物体,然后通过草图检索达到认识物体或购买物体的目的。在过去的几十年中,研究人员探索了草图在各个领域的应用场景,比如草图识别^[1-2]、基于草图的图像检索^[3-4]、基于草图的 3D 模型检索^[5]等。然而,识别非专业人员绘制的草图是一项对人类来说很有挑战性的工作。其原因有以下几点:首先,草图具有高度抽象性,例如,我们可以使用简单的几个类似火柴棍的笔画表示人类的身体;其次,我们可以使用不同的绘制方法绘制草图,所以即使是相同的物体,也有很多不同风格的表示方法,例如牛的草图,我们可以简单地绘制出牛的轮廓外形,也可以绘制出牛的花纹等细节特征;最后,草图缺少视觉信息,因为简单线条的组合就足以构成完整的草图。草图的以上 3 点特征显示了它与传统图像的差别,但是目前用于草图特征提取和识别的方法普遍将草图视作传统的纹理图像,而忽略了草图的特性。

在提取草图特征时,研究人员通常使用在传统图像领域取得很好效果的特征,比如方向梯度直方图(Histogram of Oriented Gradient, HOG)^[6]、尺度不变特征转换(Scale-Inva-

riant Feature Transform, SIFT)^[7]等。这些特征都高度依赖于丰富纹理特征的梯度,但是草图只有很简单的线条,并不具有丰富的纹理信息。神经网络在许多领域获得了巨大的成功,因此有部分工作将神经网络应用于草图的识别,但是大部分网络的结构均是针对传统纹理图像,与前文提到的特征一样,这些网络并没有考虑草图独有的稀疏性。因此,文献[8]中提出了一种专门用于草图特征提取和识别的神经网络,实验结果显示该方法的识别准确率明显高于使用其他手工特征或者神经网络结构的方法。

大部分针对草图识别的工作忽略了草图的时序性,它们将草图与传统图像视作相同的事物。但是,与传统图像一次性获得所有的像素点相比,草图中的笔画是随着时间的增长而不断增加的,这种特性称为草图的时序性。文献[8]使用了草图的时序性,利用时序性将草图分组,并同时将其输入到具有同样结构的深度网络中,但忽略了这些分组在时间序列上的相关性。

综上,本文首先根据草图的时序性将草图笔画分组,然后通过循环神经网络获取不同组之间在时间序列上的关联性,最后将循环神经网络不同时间的输出连接起来作为草图的特征,使用联合贝叶斯完成草图的识别分类工作。本文算法的整体流程如图 1 所示。

本文受国家自然科学基金(61402077)资助。

于美玉(1992—),女,硕士生,主要研究方向为计算机视觉;吴昊(1993—),男,硕士生,主要研究方向为计算机视觉;郭晓燕(1993—),女,硕士生,主要研究方向为计算机视觉;贾棋(1983—),女,博士,副教授,CCF 会员,主要研究方向为计算机视觉,E-mail:jiaqi@dlut.edu.cn;郭禾(1955—),男,硕士,教授,CCF 高级会员,主要研究方向为并行与分布式计算、计算机视觉。

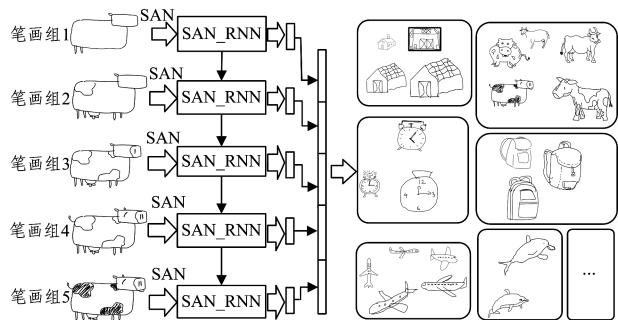


图1 算法流程图

2 相关工作

本节回顾用于草图识别的特征和网络结构。

2.1 草图识别的相关特征

2012年, Eitz^[1]公开了一个包含20000张草图的数据集,这个数据集中包含了我们生活中常见的250类事物,每类有80张草图示例。其在文中提到了人类在这个数据集上的识别准确率,这个结果证明了草图识别是一项很有挑战性的工作,因为人类的识别准确率仅为73.1%。用于草图识别的特征大致可以分为两类,一类是手工设计的特征,另一类是从神经网络中学习到的特征。

手工设计的特征一般都是在传统图像识别领域取得很好分类和识别效果的特征,比如HOG特征、SIFT特征。因此使用这些特征识别草图的一个前提是把草图当作传统的纹理图像。文献[1]划分多个图像窗口,然后在每个图像窗口中提取草图的特征,该方法也将草图没有连续的梯度和稀疏性考虑在内。文献[2]使用费舍尔向量和空间金字塔来提取草图的特征。文献[9]使用多核学习的方法学习不同特征的权重,因为不同的特征之间具有互补性,所以这种方法很大程度上提高了草图识别的准确率。

近年来,神经网络在多个领域都获得了很好的成果^[10],比如图像分类、自然语言处理等领域,这是因为深度神经网络可以使用学习的方法学习到更有区分性的层次化特征。文献[8]充分考虑了草图的独有属性,设计了专门用于草图识别的神经网络,实验结果显示该方法超越了人类创造的最高识别准确率。文献[11]提出了一种手绘草图的参数化方法,利用草图笔速和曲率信息寻找笔画的特征点,然后将特征点拟合成分段的贝塞尔曲线。该方法使用统一的参数化模型描述草图,为进一步的应用如草图识别打下了良好的基础。文献[12]根据草图笔画的方向、曲率和速度,将草图划分成多个原语形状,然后使用原语形状之间的关系描述语义符号的特征,最后使用贝叶斯分类器完成草图的分类工作。其作者将提出的方法用于电路符号的识别,实验结果显示该方法获得了显著的效果。文献[13]提出了一种基于时序关系的草图识别算法,根据笔画速度和曲率,将草图划分成各种原语形状,然后使用各个原语形状和原语形状之间的关系构造决策树,同时完成草图的识别工作。

2.2 草图识别的相关网络结构

卷积神经网络和循环神经网络是神经网络的两个重要的分支。在过去的十年中,神经网络在多个领域均获得了巨大的成功^[10,14]。LeNet^[15]是一个非常典型的神经网络,它被用

于手写数字的识别。文献[16]使用孪生网络完成了基于草图的3D模型检索工作。然而,这些工作都将草图当作传统的纹理图像,而忽视了草图固有的稀疏性、时序性等其他属性。

为了尽可能利用草图的固有属性,文献[8]设计了一个名为Sketch-A-Net的神经网络。该网络为了适应草图的稀疏性,使用更大的卷积核和池化尺寸。草图区别于传统图像的另一重要特点是草图具有时序性,获得传统图像时,我们可以一次性获取所有像素点;但是获取草图时,我们需要不断地添加草图的像素点。文献[8]通过草图笔画分组,使用不同网络提取各个笔画组的特征。这种做法体现了草图具有时序性,但是其没有将这些时序的关系体现出来。草图不是唯一一个具有时序性的事物,许多其他事物如语音、文本等也具有时序性。在处理语音识别^[17]、文本生成^[18]等任务时,研究人员引入循环神经网络来提高准确率。循环神经网络(Recurrent Neural Network, RNN)是一种适合处理输入具有“前后”关系的网络,它可以为两个输入之间的隐藏单元建立联系,然后把前一个输出传递给后一个输入。然而,原始的RNN中存在梯度消失等问题,当输入序列很长时,RNN很难完成多层神经网络的反向传播,这就容易造成梯度消失和梯度爆炸^[19]。为了克服循环神经网络的缺点,研究人员相继提出长短期记忆网络(Long Short Term Memory, LSTM)^[19]和门控制单元(Gated Recurrent Units, GRU)^[20]。GRU是轻量级的LSTM,只需要训练更少数量的参数;而且在某些情况下,GRU的效果比LSTM更好。

3 基于时序特征的草图识别方法

本节介绍本文提出的基于时序特征的草图识别方法。首先介绍门控制单元,然后探索笔画的关系和草图特征的提取,最后介绍本文使用的网络结构及草图识别的方法。

3.1 门控制单元

GRU是基础的网络结构,我们使用它来探索草图的时序性。一个GRU网络的功能是学习如何将输入序列 $X_T = (x_1, x_2, \dots, x_T)$ 映射到输出序列 $Y_T = (y_1, y_2, \dots, y_T)$,式(1)~式(5)反映了这种映射关系。

$$r_t = \text{sigm}(W_{rx}x_t + W_{rh}h_{t-1} + b_r) \quad (1)$$

$$z_t = \text{sigm}(W_{zx}x_t + W_{zh}h_{t-1} + b_z) \quad (2)$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}x}x_t + U_t(r_t \odot h_{t-1}) + b_{\tilde{h}}) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

$$y_t = W_{hy}x_t \quad (5)$$

其中, x_t 是第 t 个输入; y_t 是第 t 个输出; h_t 是隐藏单元,由 r_t, z_t, \tilde{h}_t 共同控制。操作 \odot 表示两个向量对应元素的乘法。 W 和 U 是GRU的权重矩阵, b 是GRU的权重向量。与GRU有关的更多细节信息可以参考文献[21]。

3.2 基于GRU的时序网络结构

本节首先介绍笔画分组的方法,然后介绍提取笔画组特征的方法。

3.2.1 笔画分组

对于传统的图像,我们可以一次性获得所有的像素,但是对于草图,我们不是一次性获得草图的所有内容,而是随着时间的进行,不断向草图中添加笔画,因此草图笔画的顺序是一

个很重要的信息。经过研究发现^[1],人们在画同一个物体时,有着不同的笔画顺序。如图2所示,图中有两种飞机,被2条虚线分为3个部分,中间的部分是完整的草图,左侧部分表示每个草图的第1到第10画。从左侧部分可以看出相同的物体有着不同的笔画顺序,因此,如果直接将每个笔画的特征当作GRU的输入,则不同的输入将会使得网络不够稳定。图2中右侧的部分为根据笔画顺序划分的笔画组,第一列是草图的前20%的笔画,第二列是草图的前40%的笔画,依次类推,每个笔画组都增加20%的笔画;最后一列就是完整的草图。从这对草图中可以看出,在相同列中,笔画组的差异性很小。我们统计数据集中所有草图的笔画数目,其中有54.7%的草图由6~20画构成,同时有16%的草图的笔画数少于5。因此为了减少笔画顺序带来的影响,每个草图的笔画都根据时序分成5个笔画组。如果笔画组的数目过少,则笔画组将会

既包含外轮廓信息也包含细节信息;而当笔画组数目过多时,笔画分组将会演变成一组只有1画的情况。假设草图 S 有 N 个笔画,记作 $(s_1, s_2, \dots, s_{N-1}, s_N)$,第一个笔画组包含的笔画为 $s_1 \sim s_{N/5}$,第二个笔画组包含的笔画为 $s_1 \sim s_{2N/5}$,依次类推,最后一个笔画组包含草图 S 的所有笔画。在笔画分组过程中,每组增加的笔画数目是相同的,这是因为各个笔画的重要性相同,保证了各个笔画组信息增量的公平性。为了实现鲁棒性,我们裁剪了每个笔画组的5个部分,然后又取这5个部分的镜面映射^[10]。完成这个操作后,每个笔画组对应10个草图,因此一个完整的草图对应50个草图,记作 $S = (ss_1, ss_2, \dots, ss_{49}, ss_{50})$ 。图3展示了一个笔画组中的草图输入顺序,当 t 是奇数时,输入是笔画组的裁剪,当 t 是偶数时,输入是笔画组镜面映射的裁剪。裁剪的顺序是左上、左下、右上、右下和中心。

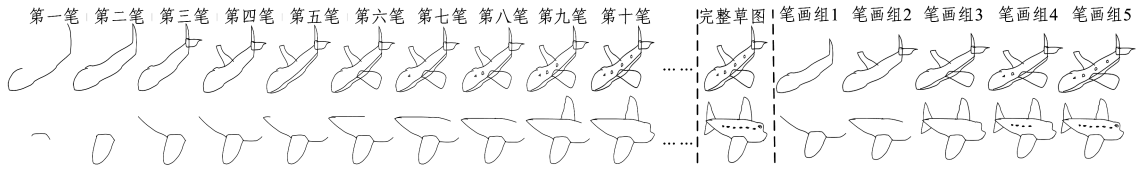


图2 笔画和笔画组对比图

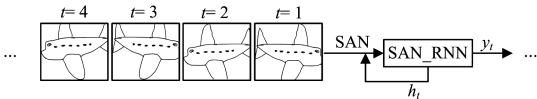


图3 同一笔画组输入顺序示意图

3.2.2 获取草图特征的方法

在提取草图特征方面,我们选择 Sketch-A-Net 提取草图的特征。与其他倾向于传统图像的网络相比,Sketch-A-Net 是研究人员专门为草图设计的网络。Sketch-A-Net 网络共有8层,前5层为卷积层,后3层为全连接层;每个卷积层对应的激活函数为 ReLU;最后一层有250个输出单元,对应 TU-Berlin 数据集^[1]中的250类。与传统图像相比,草图的纹理信息很少,为了获得更多的信息,Sketch-A-Net 网络的卷积核和池化层的尺寸更大。为了获得草图的特征,把3.2.1节中获得的草图依次输入到 Sketch-A-Net 中,然后将最后一个全连接层的512维特征当作序列中每个草图的特征。

3.2.3 多时序融合及草图分类

将3.2.2节中获得的草图特征记作 F_s ,如式(6)所示:

$$F_s = \{f_1, f_2, \dots, f_{50}\}, f_t \in \mathcal{R}^{1 \times 512}, 1 \leq t \leq 50 \quad (6)$$

根据特征的时间关系,将每个时间的特征当作RNN的输入,同时获得每个时间点的输出 y_t ,所有时间点的输出如式(7)所示:

$$Y_s = \{y_1, y_2, \dots, y_{50}\}, y_t \in \mathcal{R}^{1 \times 250}, 1 \leq t \leq 50 \quad (7)$$

不同时间点的特征描述了草图的不同特征,不同特征对草图识别的贡献度也应该不同。因此,我们选择联合贝叶斯(Joint Bayes, JB)^[22]学习不同时间点特征之间的权重。给定两个模型 $P(x_1 + x_2 | H_1)$ 和 $P(x_1 + x_2 | H_2)$,它们分别表示实例 x_1 和 x_2 属于同一类别和不同类别的概率, H_1 为两个实例属于同一类别的假设, H_2 为两个实例属于不同类别的假设。使用式(8)来描述两个实例 x_1 和 x_2 属于同一类别的程度。

$$r(x_1, x_2) = \log \frac{P(x_1 + x_2 | H_1)}{P(x_1 + x_2 | H_2)} \quad (8)$$

在本文中,草图的特征被当作 x ,使用训练GRU的数据

训练联合贝叶斯。完成学习后,计算每个测试草图和每个训练草图的相似度,最后使用K近邻(K Nearest Neighbor, KNN)完成草图的识别和分类工作。联合贝叶斯混合了不同的特征,为不同的特征分配了对应的权重,使得这些特征更具有区分性。

4 实验结果及分析

本节首先给出本文方法和现有方法在公开数据集上的实验结果及分析,然后验证联合贝叶斯的有效性。从实验结果可以看出本文方法的识别准确率比其他方法高出6%。

4.1 数据集及数据增量

TU-Berlin^[1]是一个评估草图识别算法的通用数据集,因此本文也使用这个数据集评估所提方法。这个数据集由20000张草图组成,每个类别有80张草图,所以一共有250个类别。数据集是通过 Amazon Mechanical Turk (AMT) 平台由1350名参与者完成的,因此数据集保证了物体类别的多样性和每个类别内草图风格的多样性。我们使用67%的数据进行训练,33%的数据进行测试。

在实验过程中,选择做数据增量来防止过拟合。为了增加每个类别中草图的数量,我们对每个草图做一些变换,这些变换包括水平的翻转和旋转 $[-5^\circ, -3^\circ, 0, +3^\circ, +5^\circ]$,同时也对原始草图做上、下、左、右平移 $(\pm 15^\circ)$ 。在完成数据增量后每个类别的草图数目由原来的80张增长到 $(10+8) \times 80 = 1440$ 张,所以对于250类,一共有 $1440 \times 250 = 360000$ 张草图。

4.2 实验设置

实验中,使用 Sketch-A-Net^[8]提取草图特征。Sketch-A-Net 由5个具有相同结构的网络组成,这5个网络分别用来训练5种不同尺度的草图。我们使用由尺度为 256×256 的草图训练的网络提取草图的特征,然后使用最后一个全连接层输出的512维特征作为草图的特征。使用深度学习框架 Torch^[23]实现我们的网络结构。

4.3 识别结果对比

本节展示了本文方法与其他方法的识别准确率。这些方法可以被分为两大类,第一大类是使用手工设计的特征识别草图,包括 HOG-SVM 方法^[1]、结构组合匹配方法^[24]、多核 SVM^[9]和费舍向量-空间池化方法(Fisher Vector Spatial Pooling, FV-SP)^[2];另一大类是基于神经网络的方法,包括 AlexNet^[10]、LeNet^[15]、两个版本的 Sketch-A-Net (SN1.0^[8]和 SN2.0^[25])。实验结果为各个方法相关论文中展示的结果。

表 1 列出了各种方法在 TU-Berlin 数据集上的识别准确率。从表中的结果来看,大部分基于神经网络方法的识别准确率高于使用手工设计特征的方法。使用手工设计特征方法的平均识别准确率为 63% 左右,这个识别准确率低于人类的识别准确率。这是因为这些手工设计的特征都是针对传统图像的,依赖于传统图像中丰富的纹理信息,而且传统图像是物体通过摄影设备得来的真实映射,所以这些特征并不适用于草图识别,因为草图一般都是由简单的黑色线条组成,内容稀疏,没有其他颜色纹理信息,而且草图是人类大脑对真实物体处理后的产物,具有抽象性。从表中可以看出,基于神经网络方法的平均识别准确率为 74% 左右,比人类的识别准确率高 1% 左右,比基于手工特征的方法高 11% 左右。在这些基于神经网络的方法中,SN1.0 的识别准确率为 74.90%,是第一个高于人类识别准确率的结果;而它的升级版 SN2.0 则获得了 77.95% 的准确率。基于神经网络的方法可以更好地学习草图的层次化特征,完成草图的识别工作。而本文方法的识别准确率为 84.01%,比 SN2.0 高出 6% 左右,其原因有以下几点:首先,本文充分考虑了草图的稀疏性和抽象性,所以使用专门用于识别草图的神经网络提取草图特征,而不是使用传统的用于识别图像的方法;另外,本文充分考虑了草图的时序性,将草图笔画分组,然后使用循环神经网络探索各个笔画组的先后关系,同时使用联合贝叶斯探索各个笔画组之间的重要性关系。

表 1 识别准确率对比

方法	识别准确率/%
HOG-SVM	56.00
Ensemble	61.50
MKL-SVM	65.80
FV-SP	68.90
Humans	73.10
AlexNet-SVM	67.10
AlexNet-Sketch	68.60
LeNet	55.20
SN1.0	74.90
SN2.0	77.95
本文方法	84.01

4.4 联合贝叶斯的作用

本节探索联合贝叶斯对草图识别的作用。如表 2 所列,未使用联合贝叶斯时,识别准确率只有 77.05%,使用联合贝叶斯之后,识别准确率达 84.01%,增长了大约 7%,这说明联合贝叶斯混合了不同时间点的特征后,使得草图的特征更加具有区分性,因此才能获得更高的识别准确率。如图 4 所示,我们给出了几个有无联合贝叶斯的识别结果的示例。第一列为待识别草图,第二列为无联合贝叶斯的识别结果,第三列为有联合贝叶斯的识别结果。以图中的第三行的“橱柜”为例,

它被无联合贝叶斯的方法识别为“手”,我们可以看出橱柜和手有着相似的结构特征,也就是说都有几层类矩形结构。在无联合贝叶斯时,识别草图只能根据循环神经网络输出的特征做出判断,虽然循环神经网络考虑了时序特征,但是没有考虑时序特征之间的关系,所以将“橱柜”识别成“手”;在添加联合贝叶斯之后,联合贝叶斯混合了不同的时序特征,使得草图特征更具区分性,因此给出待识别草图“橱柜”的正确分类。

表 2 有、无联合贝叶斯的识别准确率对比

方法	识别准确率/%
无联合贝叶斯	77.05
本文方法	84.01



图 4 有、无联合贝叶斯识别结果对比

结束语 本文提出了基于时序特征的草图识别方法,实验结果表明本文方法的识别准确率明显高于现有方法。所提方法充分考虑了草图的独有属性——稀疏性和时序性,同时使用循环神经网络将时序性与神经网络结合;联合贝叶斯的应用使得草图的特征更具有区分性。通过本文方法可以获得草图的特征,这些特征还可以应用于其他草图相关的领域中,例如:基于草图的图像检索,基于草图的 3D 形状检索等。

参考文献

- [1] EITZ M, HAYS J, ALEXA M. How do humans sketch object? [J]. ACM Transactions on Graphics, 2012, 31(4): 1-10.
- [2] SCHNEIDER R G, TUYTELAARS T. Sketch classification and classification-driven analysis using fisher vectors [J]. ACM Transactions on Graphics, 2014, 33(6): 174.
- [3] EITZ M, HILDEBRAND K, BOUBEKEUR T, et al. Sketch-based image retrieval: Benchmark and bag-of-features descriptors [J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(11): 1624-1636.
- [4] HU R, COLLOMOSSE J. A performance evaluation of gradient field hog descriptor for sketch based image retrieval [J]. Computer Vision and Image Understanding, 2013, 117(7): 790-806.
- [5] WANG F, KANG L, LI Y. Sketch-based 3d shape retrieval

- using convolutional neural network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA; IEEE Press, 2015; 1875-1883.
- [6] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA; IEEE Press, 2005, 1: 886-893.
- [7] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] YU Q, YANG Y, SONG Y Z, et al. Sketch-a-net that beats humans[C]// British Machine Vision Conference, BMVC 2015. Swansea, UK; BMVA Press, 2015; 1-12.
- [9] LI Y, HOSPEDALES T M, SONG Y Z, et al. Free-hand sketch recognition by multi-kernel feature learning[J]. Computer Vision and Image Understanding, 2015, 137: 1-11.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image net classification with deep convolutional neural networks[C]// Advances in Neural Information Processing Systems. Lake Tahoe, Nevada, USA; IEEE Press, 2012; 1097-1105.
- [11] 王卫, 尹建峰, 孙正兴. 一种手绘草图的快速参数化方法[J]. 计算机学报, 2006, 33(1): 264-268.
- [12] 袁贞明, 金贵朝, 张佳. 基于贝叶斯网络的在线草图识别算法[J]. 计算机工程, 2010, 36(5): 32-34.
- [13] 尹建峰, 孙正兴. 基于时序的多笔划草图识别[J/OL]. 中国科技论文在线, <http://www.paper.edu.cn/search/simple?searchType=&searchContent=%25E5%259F%25BA%25E4%25BA%258E%25E6%2597%25B6%25E5%25BA%258F%25E7%259A%2584%25E5%25A4%259A%25E7%25AC%2594%25E5%2588%2592%25E8%258D%2589%25E5%259B%25BE%25E8%25AF%2586%25E5%2588%25AB&searchDate=2003-2018&searchPage=1&searchSubject=%25E5%2585%25A8%25E9%2583%25A8&searchSort=relevant>.
- [14] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXivpreprint arXiv: 1409.1556, 2014.
- [15] LECUN Y, BOSER B E, DENKER J S, et al. Handwritten digit recognition with a back-propagation network[C]// Advances in Neural Information Processing Systems. Denver, Colorado, USA; Morgan Kaufmann, 1990; 396-404.
- [16] WANG F, KANG L, LI Y. Sketch-based 3d shape retrieval using convolutional neural networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA; IEEE Computer Society, 2015; 1875-1883.
- [17] VINYALS O, RAVURI S V, POVEY D. Revisiting recurrent neural networks for robust ASR[C]// 2012 IEEE International Conference on Acoustics, Speech and Signal Processing. Kyoto, Japan; IEEE Press, 2012; 4085-4088.
- [18] SUTSKEVER I, MARTENS J, HINTON G E. Generating text with recurrent neural networks[C]// Proceedings of the 28th International Conference on Machine Learning. Bellevue, Washington, USA; MLR. org, 2011; 1017-1024.
- [19] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [20] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014. Doha, Qatar; Association for Computational Linguistics, 2014; 1724-1734.
- [21] CHUNG J, GULCEHRE C, CHO K, et al. Gated feedback recurrent neural networks[C]// International Conference on Machine Learning. Lille, France; MLR. org, 2015; 2067-2075.
- [22] CHEN D, CAO X, WANG L, et al. Bayesian face revisited: A joint formulation[C]// 12th European Conference on Computer Vision. Florence, Italy; Springer, 2012; 566-579.
- [23] COLLOBERT R, BENGIO S, MARIËTHOZ J. Torch; a modular machine learning software library[R]. Idiap Research Report, 2002.
- [24] LI Y, SONG Y Z, GONG S. Sketch Recognition by Ensemble Matching of Structured Features[C]// British Machine Vision Conference, BMVC 2013. Bristol, UK; BMVA Press, 2013; 2.
- [25] YU Q, YANG Y, LIU F, et al. Sketch-a-net: A deep neural network that beats humans[J]. International Journal of Computer Vision, 2017, 122(3): 411-425.

(上接第 179 页)

参考文献

- [1] XU Z, WANG J, LIU Y, et al. A review of computer vision development and trends[J]. Journal of Xi'an University of Posts and Telecommunications, 2013, 17(6): 1-8.
- [2] 丁建文, 业宁, 王厚立, 等. 运用信息技术提高木材利用率和使用价值[J]. 木材加工机械, 2007, 18(1): 44-46.
- [3] ASILTURK I, UNUVAR A. Intelligent adaptive control and monitoring of band sawing using a neural-fuzzy system[J]. Journal of Materials Processing Technology, 2009, 209(5): 2302-2313.
- [4] 梅振荣, 任洪娥, 朱蒙. 基于非线性最小二乘原理的原木端面识别算法[J]. 计算机工程与应用, 2012(2): 177-178.
- [5] 刘明媚. 基于区域显著性的彩色图像分割[J]. 电子设计工程, 2013, 21(18): 133-135.
- [6] 刘越, 彭宏京, 钱素静. 基于核空间 LLE 的彩色图像分割方法[J]. 计算机学报, 2013, 40(S1): 180-183.
- [7] 梅振荣. 基于端面图像处理的原木径级识别算法研究与实现[D]. 哈尔滨: 东北林业大学, 2011.
- [8] 赵亚凤, 任洪娥. 遗传算法和同态滤波在原木端面图像处理中的应用[J]. 东北林业大学学报, 2014, 42(2): 129-132.
- [9] CHENG H D, JIANG X H, SUN Y, et al. Color image segmentation: advances and prospects[J]. Pattern Recognition, 2001, 34(12): 2259-2281.
- [10] 景林, 林耀海, 温永仙, 等. 结合色彩特征和空域特征的成捆原木轮廓识别[J]. 计算机系统应用, 2013, 22(7): 196-199.
- [11] 杨璟, 朱雷. 基于 RGB 颜色空间的彩色图像分割方法[J]. 计算机与现代化, 2010(8): 147-149.
- [12] 刘明媚. 基于区域显著性的彩色图像分割[J]. 电子设计工程, 2013, 21(18): 133-135.
- [13] CANNY J. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986(6): 679-698.
- [14] 曾俊. 图像边缘检测技术及其应用研究[D]. 武汉: 华中科技大学图书馆, 2011.
- [15] https://en.wikipedia.org/wiki/HSL_and_HSV#/media/File:HSV_color_solid_cone_chroma_gray.png.