

基于多维局部二值模式和 XGBoost 的轻量谱线删除法

吴天龙, 李强, 关欣*

天津大学微电子学院, 天津 300072

摘要 由于在手写乐谱中搜索谱线位置比较困难, 为提高乐谱谱线删除算法的稳健性, 提出了一种基于多维局部二值模式识别和 XGBoost 模型的手写乐谱谱线删除方法。根据乐谱图像的特点, 设计并改进局部二值模式算子, 提取乐谱图像中的多维局部二值模式特征算子, 组成高维特征向量, 再选择最优的 XGBoost 模型来识别乐谱谱线位置, 进而删除谱线。研究表明, 该方法在测试数据上的 F -measure 为 97.19%, 说明其具有很高的准确率和召回率; 而在三个不同子测试集上的 F -measure 分别为 96.43%, 98.36% 和 96.79%, 说明其具有很好的稳健性。相比已有的轻量谱线删除算法, 该方法的 F -measure 有所提升。

关键词 数字图像处理; 局部二值模式; 光学乐谱识别; 谱线删除

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.061006

Lightweight Staff Removal Method Based on Multidimensional Local Binary Pattern and XGBoost

Wu Tianlong, Li Qiang, Guan Xin*

School of Microelectronics, Tianjin University, Tianjin 300072, China

Abstract It is difficult to search the spectrum line in handwritten music spectrum, so in order to improve the robustness of the handwritten music spectral line deletion algorithm, a method based on multidimensional local binary pattern recognition and XGBoost model is proposed. The local binary pattern operator is designed and improved based on the characteristics of music score image, and from which the multidimension local binary pattern feature operator is extracted. Therefore, a high-dimensional feature vector is formed and the optimal XGBoost model is selected to identify the music spectral line location, then the line is deleted. The research results show that F -measure of this method is 97.19% on the test data, which illustrates that the method has a high accuracy and recall rate. F -measure is 96.43%, 98.36% and 96.79% respectively on three different test subsets, which shows that it has good robustness. Compared with existing lightweight spectrum line deletion algorithm, the F -measure of this method is relatively improved.

Key words digital image processing; local binary pattern; optical music recognition; spectrum line deletion

OCIS codes 100.2000; 100.2960; 100.3008

1 引言

目前, 在世界各地的图书馆和博物馆中保存着大量纸质乐谱, 而这些珍贵的纸质乐谱资料随时会面临丢失和损坏的危险, 由于纸质载体不易传播和交流, 所以人们将纸质乐谱转换为计算机可直接识

别和利用的符号形式(MIDI 或 XML)来解决以上难题。光学乐谱识别(OMR)^[1]技术由此产生, 可实现纸质乐谱向符号乐谱的自动转换。一般来说, OMR 技术首先要删除谱线, 再进行音符识别。尽管少数学者提出可在不删除谱线的情况下进行音符识别, 但是在乐谱图像中, 谱线和音符会重叠交织在

收稿日期: 2018-09-26; 修回日期: 2018-10-17; 录用日期: 2018-10-25

基金项目: 国家自然科学基金(61471263)、天津市自然科学基金(16JCYDJC31100)

* E-mail: guanxin@tju.edu.cn

一起,谱线的存在会给后面的音符识别造成较大困难,因此有效的 OMR 技术需要先进行谱线删除。在理想的情况下,谱线是一组平行的细线,相对来说删除谱线较容易,但对于手写乐谱来说,要做到精确删除并不容易。而且受音乐家手写乐谱风格的影响,在手写乐谱图像中音符的形式更加复杂多变,因此区分谱线和音符也变得更为困难。目前,谱线删除方法主要分为两类:基于结构特征判别的方法和基于机器学习分类的方法。

基于结构特征判别的方法是根据谱线的水平结构特点进行搜索删除。Fujinaga 等^[2]提出了游程编码的谱线检测方法。Szwoch 等^[3]采用水平投影的方法。Miyao 等^[4]采用动态规划的方法进行匹配。Bainbridge 等^[5]使用直线追踪的方法,但在谱线间断处的效果较差。Roach 等^[6]提出将像素转化成向量,再根据向量的方向和长度进行谱线的删除。Carter 等^[7]采用行邻接图的方法。Ng^[8]使用骨架化方法。上述这些方法在 Dalitz 的一篇文献^[9]里做了详细的对比分析。dos Santos Cardoso 等^[10]根据图论的方法,由乐谱的左侧边缘开始,找出一条到乐谱右侧边缘的最短路径,并采用动态规划的方法优化最短路径。Su 等^[11]首先估计谱线宽度和间距,得到大致的谱线位置和形状,然后在原始乐谱图像中进行谱线删除并得到最终结果。Géraud 等^[12]提出了一种基于形态学的方法。基于结构特征判别的方法删除谱线速度较快,并能较好地删除理想乐谱谱线,但对于手写乐谱中谱线不规则的情况,该方法搜索谱线会变得困难,因此算法的准确性和稳定性有所下降。

基于机器学习分类的方法是将谱线检测问题归结为对乐谱图像中的黑色像素进行二分类的问题。首先根据乐谱图像的特点进行手动选取并计算像素点的特征,也可通过学习得到特征,然后利用已标注的谱线和音符标签的图像数据来训练模型。训练好的模型可将待处理图像像素点区分为谱线像素和音符像素。Calvo-Zaragoza 等^[13]提出了一种基于邻域特征进行像素分类的方法,通过实验选取最佳邻域窗口,但该方法未能结合乐谱图像的特点设计窗口。Montagner 等^[14]使用图像算子学习的方法来删除谱线,对于数据量较小的数据集,可使用先验的自定义窗口算子,对于更复杂的删除任务,可通过学习来获得窗口算子。该算法稳健性较好,但不适用于谱线间断或存在一定旋转角度的乐谱图像。

近年来,一些学者提出采用深度学习的方法进

行谱线删除,而深度学习是直接利用图像块数据,不需要显示计算特征,便可直接用于图像的语义分割^[15]。Calvo-Zaragoza 等^[16]采用卷积神经网络(CNN)方法,详细对比并分析了微调如窗口大小等参数对谱线删除效果的影响。Aguilar 等^[17]采用 CNN 和图像算子相结合的方法。Gallego 等^[18]采用了自编码器方法。在包含 3D 扭曲和局部噪声的非理想情况的 GREC/ICDAR 2013 手写乐谱谱线删除竞赛数据集^[19]上,深度学习谱线删除方法取得了较好的结果。然而,训练深度学习模型需要大量数据,并且训练阶段耗时较长,在轻量计算的 OMR 技术应用场景下较难采用深度学习的方法。目前虽然有一些针对移动设备的深度学习架构,但是该设备处理器的计算能力尚不能满足深度学习本地化的计算需求。针对存在扭曲和噪声的复杂手写乐谱图像,本文提出了一种谱线删除的轻量计算方案。局部二值模式(LBP)可以考虑乐谱图像的纹理特点,并且能充分体现谱线和音符纹理的特征。该表征图像纹理信息的方法,可用于人脸识别^[20]和图像分类^[21]等领域。

在机器学习谱线删除方案中可利用的特征方面,Meng 等^[22]曾提出采用 LBP 算子描述谱线和音符的特征,但对于非理想情况下的复杂手写乐谱图像,单一的 LBP 算子区分谱线和音符的能力不佳,因此,本文提出了多维邻域信息的 LBP 算子集合作为分类时采用的特征向量。为了区分谱线和音符的二分类模型,选择了 Chen 等^[23]在 2016 年提出的 XGBoost 算法,这是一种基于提升决策树的学习算法。从模型角度来看,XGBoost 算法引入正则项,并控制模型的复杂度,从而有效地控制了过拟合;从优化计算角度来看,XGBoost 算法利用二阶导数信息,是寻找分裂节点的近似算法,可加快迭代收敛的速度,并减小内存消耗。此外,XGBoost 算法还支持并行计算。综合考虑上述特点,发现其适合轻量计算场合。

2 多维 LBP 特征向量

LBP 最早由 Ojala 等^[24]提出,是描述图像局部纹理特征的算子。基本 LBP 算子只采用一种尺寸窗口提取特征算子,因此缺少其他尺度的纹理信息。对于二值化后的乐谱图像,某像素点是音符还是谱线均与该像素点的局部纹理特征密切相关。在理想乐谱图像中,谱线具有明显的水平方向性,而大部分音符具有明显的垂直方向性,同时音符所占像素宽

度一般大于谱线的宽度,综合以上特点可构造出水平和垂直的 LBP 特征算子。为了考虑非理想的情况和手写音符的多样性,又设置了与水平方向角度呈 45°的 LBP 特征算子。

2.1 谱线检测多方向 LBP 算子

在所选矩形窗口的范围内,由中心向外,根据各邻域像素的特征可构造第一个 LBP 算子(L_1),并将该算子进行顺时针旋转,每隔 45°作为下一个 LBP 算子。这样在一个固定矩形窗口内,共可构造 8 个 LBP 算子,作为候选特征,图 1 所示为 9×9 窗口内的 8 个候选 LBP 特征算子。

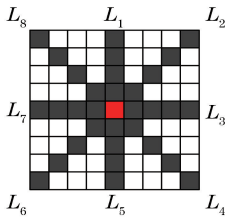


图 1 候选 LBP 特征算子

Fig. 1 Candidate LBP feature operators

图 2 为实验所用的乐谱图像是二值图像,音符和谱线构成的前景色像素灰度值为 1,背景色像素灰度值为 0。

对于 $n \times n$ 的矩形窗口,每个候选 LBP 特征算子包含 $(n-1)/2$ 个相邻像素,再根据基本 LBP 特征算子的编码规则在矩形窗口内由中心向外进行二进制编码,可得到每个候选 LBP 特征算子的二进制编码,编码规则可表式为

$$s_i^j = \begin{cases} 1 & g_i \geq g_c, i = 1, 2, \dots, \frac{n-1}{2} \\ 0 & g_i < g_c \end{cases}, \quad (1)$$

式中 g_c 为中心点像素值, g_i 为相邻像素值, s_i^j 为



图 2 实验所用乐谱示例。(a)输入乐谱;(b)真值标签

Fig. 2 Example of score used in experiment.

(a) Input score; (b) ground-truth label

第 j 个 LBP 特征算子的第 i 个相邻像素所对应的二进制编码。

然后分别计算并得到每个候选 LBP 特征算子的 LBP 值,由中心向外第 i 个像素赋予 2^{i-1} 权重,则第 j 个 LBP 特征算子的 LBP 值 L_j 可表示为

$$L_j = \sum_{i=1}^{(n-1)/2} s_i^j \times 2^{i-1}, j = 1, 2, \dots, 8, \quad (2)$$

如图 3 所示,以 9×9 窗口为例,给出了 8 个候选 LBP 特征算子的计算过程并对邻域像素编码,编码的顺序由中心向外依次进行。以图 3 中的 L_1 为例,根据编码规则可得 4 位二进制编码(1111),然后由中心向外第 i 个像素赋予 2^{i-1} ,将其转换为十进制数: $L_1 = 1 \times 2^0 + 1 \times 2^1 + 1 \times 2^2 + 1 \times 2^3 = 15$,其余 LBP 特征算子计算方法以此类推。其他尺寸窗口的 LBP 特征算子计算过程与此规则相同,只是编码位数和对应的权重根据尺寸有所变化。

2.2 不同尺寸谱线检测 LBP 算子

对于平均谱线宽度为 m 个像素的乐谱图像,将选择 $(2m+1) \times (2m+1)$ 大小的窗口作为最小 LBP

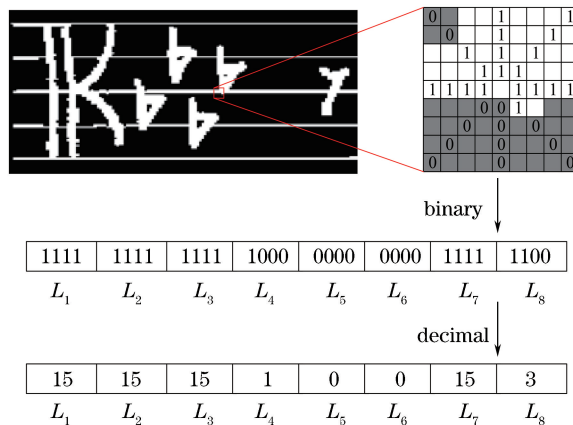


图 3 LBP 特征算子提取过程

Fig. 3 Process of LBP feature operator extraction

的特征窗口,若窗口再小,则 LBP 特征算子不能提供区分谱线和音符的有效信息。由于 LBP 特征算子需选取中心像素点,所以 LBP 窗口尺寸要选择奇数值,使得相对方向的 LBP 特征算子编码位数相同。由 LBP 特征值的计算公式可知,随着窗口增大,单个 LBP 特征值将呈指数形式增长,为了加快特征计算速度,应尽量选择较小的窗口。综上所述,候选窗口为 $(2m + 1) \times (2m + 1)$ 到 $(2m + 11) \times (2m + 11)$ 的奇数尺寸窗口,并增加了组合窗口,窗口尺寸将通过实验确定。

3 谱线检测与删除系统

首先,根据数据集中谱线的宽度确定候选窗口尺寸,再对前景色像素点提取所有候选窗口的 LBP

特征算子,并构成高维特征向量。然后,在 XGBoost 算法中选择适合谱线识别的目标函数和正则项,并利用训练数据进行监督学习,得到最优 XGBoost 模型。图 4 所示为基于多维 LBP 特征算子和 XGBoost 算法的手写乐谱谱线删除算法的测试阶段流程图。

3.1 谱线识别算法选择

XGBoost 算法采用了梯度提升决策树 (GBDT) 的思想,以决策树为基函数,而 XGBoost 模型由多个决策树组成,后面的决策树将拟合前面预测的残差,最终得到的预测值是所有决策树预测结果的和。其模型可表示为

$$y_r = \sum_{k=1}^K f_k(x_r), f_k \in F, \quad (3)$$

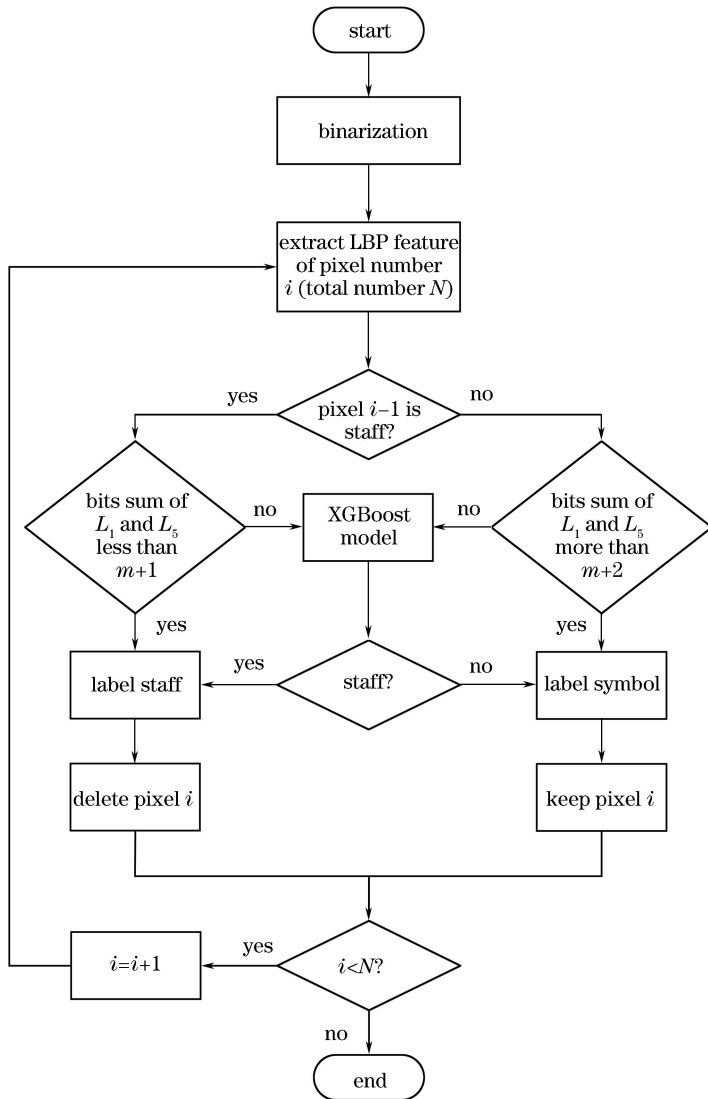


图 4 测试阶段流程图

Fig. 4 Flow chart of test stage

式中 y_r 为模型的第 r 个样本预测值, f_k 为第 k 个决策树基函数, K 为决策树基函数的总数量, x_r 为第 r 个输入样本, F 为假设空间。

对于每一棵决策树, 首先要确定目标函数。目标函数 $L^{(t)}$ 可表示为

$$L^{(t)} =$$

$$\sum_{r=1}^M l\{[Y_r, y_r^{(t-1)} + f_t(x_r)]\} + \Omega(f_t), \quad (4)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2, \quad (5)$$

式中 $l(\cdot)$ 为损失函数, Y_r 为真值, M 为样本总数量, $y_r^{(t-1)}$ 为第 $t-1$ 轮迭代的预测值, f_t 为第 t 轮迭代的预测值, $\Omega(\cdot)$ 为正则项, T 为决策树叶节点的个数, ω 为叶节点的输出值, γ 和 λ 为对应的系数。

谱线删除属于二分类任务, 因此损失函数的选择是对数损失函数, 对数损失函数 $l(Y, y)$ 可表示为

$$l(Y, y) = -Y \ln(y) - (1 - Y) \ln(1 - y), \quad (6)$$

式中 Y 为真值, y 为预测值。

通过 LBP 特征算子提取训练样本数据, 得到其特征维度较高, 同时单个 LBP 特征算子的取值数量较多, 采用决策树模型进行特征选择和分裂节点选择时, 计算复杂, 而 XGBoost 算法支持并行计算, 可大幅度提高计算效率, 也可自定义损失函数, 并从中选择合适的对数损失函数, 因此更适合二分类的谱线删除任务。同时, XGBoost 算法包含多个控制模型复杂程度的超参数, 可更加细致地优化模型。

3.2 谱线检测加速处理模块

在对任意乐谱图像的谱线进行检测和删除时, 如果待处理图像是灰度图像, 可以采用阈值法进行二值化处理, 并将待检测乐谱图像前景色像素点利用 LBP 特征算子进行提取。有两种预测的可能: 若相邻前一像素点被预测为谱线, 并且当前像素点的 L_1 和 L_5 特征的二进制编码的所有位数加和小于 $m+1$ (m 为平均谱线宽度), 则当前像素点可直接判断为谱线并删除; 若相邻前一像素点被预测为音符, 并且当前像素点的 L_1 和 L_5 特征的二进制编码的所有位数加和大于 $m+2$, 则当前像素点直接判断为音符。再利用训练好的 XGBoost 模型进行预测, 删除预测为谱线的像素点, 遍历整幅图像中的所有前景色像素点, 该方法可加快算法的速度。

4 谱线检测与删除实验

实验所用数据集可分为训练集和测试集。

XGBoost 模型参数可根据经验设置网格并划分搜索范围, 再利用带标签乐谱图像训练模型, 利用测试集评估性能。实验所用计算机配置为英特尔至强 E3-1270 v6 3.8 GHz 处理器。

4.1 实验用数据集

实验采用了 GREC/ICDAR 2013 手写乐谱谱线删除竞赛数据集^[19], 数据集可分为二值图像和灰度图像, 在本文中仅使用二值图像, 其中包括 4000 幅训练乐谱图像和 2000 幅测试乐谱图像, 每一张乐谱都有对应的真值标签, 即删除谱线后的图像。测试乐谱图像分为 3 个子集: 500 幅 3D 扭曲图像 TS1, 500 幅局部噪声 TS2 及 1000 幅叠加以上两种退化操作 TS3。

由于本文算法需逐个提取像素并训练样本, 同时每幅乐谱图像约包含 200000 个前景色像素点, 所以使用全部训练乐谱图像会导致训练样本过大, 进而使训练时间较长, 而且像素集训练样本会存在大量相同数据, 所以在保证训练样本充分的前提下, 尽量减少训练样本数目, 因此在 4000 幅训练乐谱图像中随机选择 40 幅即可, 同时每一幅乐谱图像随机选择一个 400×400 的乐谱图像块, 以此来训练 XGBoost 模型, 其中测试乐谱图像用来评估模型。

4.2 系统性能评价指标

为了定量评估方法性能, 采用了 F -measure 作为评价指标, 可表示为

$$F = \frac{2 \times R_{TP}}{2 \times R_{TP} + R_{FP} + R_{FN}} \times 100\%, \quad (7)$$

式中 R_{TP} 为将音符预测为音符的数目, R_{FP} 为将谱线预测为音符的数目, R_{FN} 为将音符预测为谱线的数目。

4.3 LBP 窗口选择

如前文所述, LBP 窗口的尺寸需通过实验进行选择。数据集的平均谱线宽度为 3 个像素, 因此实验使用候选窗口尺寸为 7×7 到 17×17 的奇数维窗口, 并增加了部分组合窗口。XGBoost 算法使用第三方库, 从中选择默认参数。表 1 列出不同窗口下的 F -measure, 由于 2 种窗口组合数量较多, 所以只列出实验结果较好的部分组合窗口的结果。实验结果表明, 3 种及 3 种以上窗口组合并不优于单窗口和 2 种窗口组合, 同时运行速度更慢。

对于不同尺寸的窗口, 该算法都有很高的 F -measure, 说明本谱线识别方法具有很强的稳定性。其中, 对于 TS1, 13×13 窗口表现最好。对于 TS2 和 TS3, 9×9 和 13×13 的组合窗口可得到最好结

表1 不同尺寸窗口的 F -measure

Size	TS1	TS2	TS3	Whole
7×7	94.13	97.93	95.76	95.94
9×9	96.02	98.12	96.57	96.90
11×11	95.88	98.06	96.52	96.82
13×13	96.26	98.15	96.52	96.97
15×15	95.98	98.11	96.43	96.84
17×17	94.97	98.08	95.85	96.30
7×7+9×9	95.98	98.06	96.21	96.75
9×9+11×11	95.92	98.10	96.38	96.80
11×11+13×13	96.03	98.07	96.25	96.78
9×9+13×13	96.16	98.34	96.65	97.05
13×13+17×17	96.09	98.22	96.17	96.82

果,并且其全局(Whole)平均 F -measure 也是最优的。因此,对于后续实验,窗口选择为 9×9 和 13×13 的组合窗口,特征总数为 16 维。

4.4 分类器性能对比及分析

对比不同的分类器如支持向量机(SVM)、自适应提升(Adaboost)、K-近邻(KNN)和随机森林(RF)的性能。SVM 通过最大化支持向量间隔来学习分类超平面,可使用核函数完成线性不可分任务,本实验采用的是高斯核函数。Adaboost 采用集成学习思想,在训练过程中增加错误分类数据的权重,不断迭代直到满足停止条件,最终分类器是每次迭代所产生的弱分类器的加权组合。KNN 通过寻找与目标样例距离最近的 K 个数据点,然后统计这 K 个点的标签数,选取具有最多数据的标签作为最终的分类结果。RF 也是一种集成学习方法,通过许多弱分类的组合得到最终分类决策模型,RF 与 Adaboost 和 XGBoost 不同的是,分类器之间没有依赖关系,最终分类结果取决于所有弱分类器的投票表决,多数取胜。以上 4 种分类器均采用 Scikit-learn 机器学习库,XGBoost 使用第三方库,通过网络搜索得到最优参数。

表 2 是不同分类器的对比结果,对于 3 个子测试集,使用 XGBoost 算法的 F -measure 均是最高的。与 SVM 和 KNN 相比,XGBoost 算法是一种集成学习方法,有更强大的分类性能。提取特征算子后的训练样本具有明显的不均衡性,采用 RF 或 Adaboost 会使 F -measure 精度有所下降,而 XGBoost 算法能更好地应对样本不均衡的数据。

本文使用 XGBoost 算法作为最终分类器,图 5 所示是通过最优 XGBoost 模型处理的一个乐谱片段,对比结果和真值标签可以看出,实验结果很好地

表2 不同分类器的 F -measure 对比

Classifier	TS1	TS2	TS3	Whole
SVM	95.56	96.27	96.45	96.09
Adaboost	95.31	96.10	96.33	95.91
KNN	95.89	97.02	96.20	96.37
RF	96.11	97.54	96.36	96.67
XGBoost	96.43	98.36	96.79	97.19

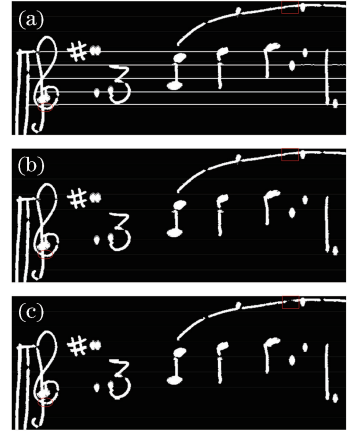


图5 乐谱片段。(a)输入乐谱;(b)真值标签;(c)处理结果
Fig. 5 Results of score part. (a) Input score; (b) ground-truth label; (c) processing result

保留了音符的完整性,但是在线宽较小的音符上存在过删除现象(如图 5 矩形框部分),这样的音符像素点 LBP 特征算子会和多数谱线像素点 LBP 特征算子相同,同样在线宽较大的谱线上也会存在欠删除现象(如图 5 椭圆框部分),这样的谱线像素点 LBP 特征算子也会和多数音符像素点 LBP 特征算子相同,因为 LBP 特征算子只提取了所选窗口内的局部纹理信息,对于具有相同 LBP 特征算子的音符像素点和谱线像素点,XGBoost 算法将不能正确分类。

4.5 谱线检测性能对比及分析

将实验结果与 CVC2013 竞赛参与者所提供的方案^[19]结果进行比较,包括以色列特拉维夫大学(TAU)方法、新加坡国立大学(NUS)方法、下莱茵应用科学大学 iPattern 研究所-直线检测(NUASi-lin)方法、下莱茵应用科学大学 iPattern 研究所-骨架化(NUASi-skel)方法、EPITA 研发实验室(LRDE)方法、计算机工程与系统研究所(INESC)方法。表 3 列出了所提方法与其他方法对比结果(F -measure),可以看出,对于 TS2,所提方法的 F -measure 值明显高于其他算法,对于 TS1 和 TS3,所提方法的 F -measure 值略低于 LRDE 而高于

表3 不同算法的 F -measure 对比Table 3 F -measure comparison of different methods %

Method	TS1	TS2	TS3	Whole
TAU	85.72	81.72	82.29	83.01
NUS	69.85	96.25	67.43	75.24
NUASI-lin	94.99	94.86	94.00	94.29
NUASI-skel	94.25	93.80	92.92	93.34
LRDE	97.73	96.86	96.98	97.14
INESC	89.29	97.72	88.52	91.01
Proposed approach	96.43	98.36	96.79	97.19

其他算法。TS1 和 TS3 测试子集都叠加了 3D 扭曲,而 TS2 只叠加了局部噪声,说明本文算法不能很好地处理 3D 扭曲乐谱图像,因为 3D 扭曲乐谱图像扭曲位置会有多段较宽的谱线,这样的谱线像素点具有的 LBP 特征算子更接近音符像素点所具有的 LBP 特征算子,所以对模型会出现较多的误分类,但是在全测试集上,所提方法具有最优的结果。在 3 个子测试集上,所提方法均具有很好的效果,可见所提方法具有很高的稳定性。

5 结 论

提出了一种结合多维 LBP 和 XGBoost 算法的手写乐谱谱线删除方法。针对乐谱图像的特点,提出一种多维 LBP 特征算子。文中提出的轻量方法可监督学习,仅使用了全部训练乐谱图像的 0.05% 作为训练数据并与其相比较,若深度学习要达到同样的谱线识别性能,通常需采用全部训练数据集进行 XGBoost 模型训练。实验结果表明所提 LBP 特征算子能较好地表征谱线和音符之间的差异,使用较少的训练数据可使 XGBoost 模型达到良好的识别效果。然而本文提出的 LBP 特征算子是对乐谱图像局部纹理信息的描述,尚有少量音符和谱线具有相同的 LBP 特征算子,显然在这种情况下,模型无法正确地区分音符和谱线,后续工作可进一步结合乐谱图像的结构特点改进方法。本文算法测试运行平均时间为 300 s/幅,目前无法做到实时处理,后续可考虑并行化设计,提高运行速度。实验结果证明,本文提出的谱线删除方法有较高的准确性和稳定性,相比传统算法有一定程度的提升。同时,XGBoost 模型更适合需轻量计算的移动设备,使得 OMR 系统可在移动设备中实现。

参 考 文 献

[1] Rebelo A, Fujinaga I, Paszkiewicz F, *et al.* Optical music recognition: state-of-the-art and open issues

- [J]. International Journal of Multimedia Information Retrieval, 2012, 1(3): 173-190.
- [2] Fujinaga I. Staff detection and removal [M/OL] // George S E. Visual perception of music notation: on-line and off line recognition. 2004: 1-39. [2018-09-26]. <https://www.igi-global.com/gateway/chapter/31053>.
- [3] Szwoch M. A Robust Detector for Distorted Music Staves [C] // Gagalowicz A, Philips W. Computer analysis of images and patterns. CAIP 2005. Lecture notes in computer science. Berlin, Heidelberg, Springer, 2005, 3691: 701-708.
- [4] Miyao H. Stave Extraction for printed music scores [M] // Yin H, Allinson N, Freeman R, *et al.* Intelligent data engineering and automated learning-IDEAL 2002. Lecture notes in computer science. Berlin, Heidelberg: Springer, 2002, 2412: 562-568.
- [5] Bainbridge D, Bell T C. Dealing with superimposed objects in optical music recognition [C] // 6th International Conference on Image Processing and its Applications. July 17, 1997, Dublin, Ireland: IET Digital Library, 1997: 756-760.
- [6] Roach J W, Tatem J E. Using domain knowledge in low-level visual processing to interpret handwritten music: an experiment [J]. Pattern Recognition, 1988, 21(1): 33-44.
- [7] Carter N P, Bacon R A. Automatic recognition of printed music [M] // Baird H S, Bunke H, Yamamoto K. Structured document image analysis. Berlin, Heidelberg: Springer, 1992: 456-465.
- [8] Ng K. Music manuscript tracing [C] // Blostein D, Kwon Y B. Graphics recognition algorithms and applications. GREC 2001. Lecture notes in computer science. Berlin, Heidelberg: Springer, 2001, 2390: 330-342.
- [9] Dalitz C, Droettboom M, Pranzas B, *et al.* A comparative study of staff removal algorithms [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(5): 753-766.
- [10] dos Santos Cardoso J, Capela A, Rebelo A, *et al.* Staff detection with stable paths [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(6): 1134-1139.
- [11] Su B L, Lu S J, Pal U, *et al.* An effective staff detection and removal technique for musical documents [C] // 10th IAPR International Workshop on Document Analysis Systems, March 27, 2012, Gold Coast, QLD, Australia. New York: IEEE,

- 2012: 160-164.
- [12] Géraud T. A morphological method for music score staff removal[C] // IEEE International Conference on Image Processing (ICIP), Oct. 27, 2014, Paris, France. New York: IEEE, 2014: 2599-2603
- [13] Calvo-Zaragoza J, Micó L, Oncina J. Music staff removal with supervised pixel classification [J]. International Journal on Document Analysis and Recognition, 2016, 19(3): 211-219.
- [14] Montagner I S, Hirata N S T, Hirata R. Staff removal using image operator learning [J]. Pattern Recognition, 2017, 63: 310-320.
- [15] Guo C C, Yu F Q, Chen Y. Image semantic segmentation based on convolutional neural network feature and improved superpixel matching [J]. Laser & Optoelectronics Progress, 2018, 55(8): 081005.
郭呈呈, 于凤芹, 陈莹. 基于卷积神经网络特征和改进超像素匹配的图像语义分割[J]. 激光与光电子学进展, 2018, 55(8): 081005.
- [16] Calvo-Zaragoza J, Pertusa A, Oncina J. Staff-line detection and removal using a convolutional neural network[J]. Machine Vision and Applications, 2017, 28(5/6): 665-674.
- [17] Aguilar F D J, Hirata N S T. Image operator learning coupled with CNN classification and its application to staff line removal [C] // 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Nov. 9, 2017, Kyoto, Japan. New York: IEEE, 2017: 53-58.
- [18] Gallego A J, Calvo-Zaragoza J. Staff-line removal with selectional auto-encoders [J]. Expert Systems with Applications, 2017, 89: 138-148.
- [19] Visaniy M, Kieu V C, Fornés A, *et al.* The ICDAR 2013 music scores competition: staff removal [C] // 12th International Conference on Document Analysis and Recognition (ICDAR 2013), Aug. 2013, Washington DC, USA. New York: IEEE, 2013: 1407-1411.
- [20] Yang H X, Chen Y, Zhang F, *et al.* Face recognition based on improved gradient local binary pattern [J]. Laser & Optoelectronics Progress, 2018, 55(6): 061004.
杨恢先, 陈永, 张翡, 等. 基于改进梯度局部二值模式的人脸识别[J]. 激光与光电子学进展, 2018, 55(6): 061004.
- [21] Ye Z, Bai L. Hyperspectral image classification based on principal component analysis and local binary patterns [J]. Laser & Optoelectronics Progress, 2017, 54(11): 111006.
叶珍, 白璘. 基于主成分分析与局部二值模式的高光谱图像分类[J]. 激光与光电子学进展, 2017, 54(11): 111006.
- [22] Meng F A, Li Q, Shen Y T, *et al.* Staff detection and removal based on local binary patterns [J]. Journal of Frontiers of Computer Science & Technology, 2017, 11(12): 2015-2022.
孟凡奥, 李镡, 申一汀, 等. 基于局部二进制模式的乐谱谱线检测与删除[J]. 计算机科学与探索, 2017, 11(12): 2015-2022.
- [23] Chen T, Guestrin C. XGBoost: a scalable tree boosting system [C] // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 13, 2016, San Francisco, California, USA. New York: ACM, 2016: 785-794.
- [24] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions [J]. Pattern Recognition, 1996, 29(1): 51-59.