刑事司法科学与治理
Criminal Justice Science & Governance
2020 年 第 1 卷 第 2 期

基于机器学习的线条痕迹检验研究 *

杨 敏[1,2] 牟 丽[1,2] 付一鸣[3]

1. 中南财经政法大学刑事司法学院，武汉；
2. 司法鉴定技术应用与社会治理学科创新基地，武汉；
3. 广东警官学院刑事技术系，广州

摘 要 | 传统的工具痕迹人工检验鉴定方法缺乏客观性受到法庭的强烈挑战。对此挑战，国内外主要研究了线条特征的提取与特征之间相关性的计算机处理方法。研究了基于人工智能的机器学习算法，对三种工具制作线条痕迹的 2D 图像数据集做了 4 组实验，通过提取线条痕迹的四种局部二进制模式（LBP）的衍生算子构建痕迹的特征向量，使用随机森林算法对带标签的特征向量进行监督学习。实验结果表明本方法对于在相同条件下制作的线条痕迹具有较高的识别率，且能够有效克服工具痕迹 2D 图像数据光照不稳定的缺点，也避免了现有方法中因人工预设置参数给痕迹检验带来的困难和检验结果的不确定性。

关键词 | 法庭科学；线条痕迹；人工智能；机器学习

Copyright © 2020 by author (s) and SciScan Publishing Limited
This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

一、引言

工具痕迹具有犯罪现场遗留率高、不易被破坏、易提取等特点，在确定侦查方向和范围、犯罪现场重建和诉讼证据等方面有着举足轻重的作用。工具与客体表面相互作用且做相对运动时，客体表面会留下线条痕迹。实际工作中，线条痕迹的检验鉴定是基于工具痕迹上的种类特征和细节特征，判断检材和样本特征的相似性来确定两者是否同一[1]。传统工具痕迹检验方法的有效性受到来自法庭的挑战，问题主要集中在工具痕迹检验方法缺乏客观性和非定量化[2][3]。针对该挑战，大量围绕统计学并以

* 基金项目：国家自然科学基金面上项目（No. 61471134）。

[1] Winston M D . Recruitment theory：identification of those who are likely to be successful as leaders [J] . Journal of Library Administration，2001（32）：19–33.

[2] Edwards H，Gatsonis C，Berger M，et al. Strengthening forensic science in the United States：A path forward [J] . National Academy Press，2009（13）：978.

[3] Schwartz A. Systemic Challenge to the Reliability and Admissibility of Firearms and Tool Mark Identification [J] . The Columbia Science and Technology Law Review，2005（6）：1–42.

计算机为辅助的工具痕迹检验方法逐渐成为研究热点，线条工具痕迹计算机检验是将来的发展方向[1]。

螺丝刀线条痕迹因具有一致性的连续线条而常常被选择作为研究对象。Neel 等[2]针对螺丝刀线条工具痕迹的 2D 灰度图像和 3D 形貌数据，研究了线条痕迹间的相关统计特性，验证了已知匹配和已知非匹配线条痕迹统计定量检验的有效性，指出 2D 灰度图像数据易受光照影响。Faden 等[3]针对 44 把连续制造的螺丝刀，以每把螺丝刀的两个刀口面分别以 30°、60° 和 85° 接触角在铅块上制作痕迹，然后使用 3D 轮廓仪采集得到包含线条高度、深度和宽度信息的 1D 特征向量，将向量划分为 35 个小段，分别比较两个痕迹的向量各段的最大相关值来确定两痕迹间的关系。该研究说明了数学方法用于痕迹定量化检验是可行的。Chumbley 等[4]对 Faden 提出的痕迹相关性比对算法进行了优化，把相关性计算分为优化和证实两个步骤，改进的算法能够提高痕迹检验的效果。Baiker 等[5]提出了一种线条工具痕迹自动比对方法，结合多尺度配准线条痕迹，使用全局交叉相关值作为工具痕迹的相似测度，可识别螺丝刀工具在接触角度 30° 误差内形成的线条痕迹。Louis 等[6]研究了螺丝刀的轴旋转角度对形成的工具痕迹产生的影响，研究表明痕迹间的相似性和可重复性随着旋转角度（0° ~ 75° 度间）增大而逐渐减小。同样地，Baiker 等[7][8]使用计算机仿真讨论了工具与承痕客体接触角度的变化、承痕客体材料性质和痕迹的深度对痕迹检验的影响。Nicholas 等[9]使用 9 个螺丝刀制作理想线条模式的可复制集，把线条痕迹图像沿着某个方向扫描，把图像灰度编码成 1 个高维特征向量，首次提出使用多统计（PLS-DA）、主成分分析、支持向量机（PCA-SVM）的模式识别方法对特征向量进行分类识别，其研究表明可用模式统计识别方法实现线条痕迹比对的稳定性。Bachrach 等[10]使用螺丝刀以不同接触角（15°、30° 和 45°）先后在铅块上制作线条痕迹，再以 30° 角在铅块上制作痕迹，使用舌槽钳剪切铅丝、黄铜丝和镀锌铁丝，最后使用共焦显微镜采集痕迹样本选定区域的 3D 形貌数据，对数据预处理后依据轮廓高度和深度信息构建 1D 特征向量，计算特征向量的相对距离作为相似测度来判断痕迹是否匹配。

提取线条痕迹的 3D 形貌数据，把其映射到 1D 数据，再使用交叉相关值估计 2 个痕迹的相似性，这种方法是线条痕迹比对检验研究的热点。除了使用相关性参数测量痕迹间的相似性以外，也有其他方法应用于工具痕迹的检验。Taroni 等[11]使用基

[1] Nichols R. Firearm and toolmark identification criteria: a review of the literature [J]. J. Forensic Sci, 2003 (48): 318-327.

[2] Neel M, Wells M. A Comprehensive Statistical Analysis of Striated Tool Mark Examinations Part 1: Comparing Known Matches and Known Non-Matches [J]. AFTE Journal, 2007, 39 (3): 176-198.

[3] Faden D, Kidd J, Craft J, et al. Statistical confirmation of empirical observations concerning tool mark striae [J]. AFTE J, 2007, 39 (3): 205-214.

[4] Chumbley L S. Validation of tool mark comparisons obtained using a quantitative, comparative, statistical algorithm [J]. J. Forensic Sci, 2010: 55 (4): 953-961.

[5] Baiker M I, Keereweer R, Pieterman E, et al. Quantitative comparison of striated toolmarks [J]. Forensic Sci Int, 2014 (242): 186-199.

[6] Louis Garcia D, Pieterman R, Baiker M. Influence of the axial rotation angle on tool mark striations [J]. Forensic Sci Int, 2017 (279): 203-218.

[7] Baiker M, Pieterman R, Zoon P. Tool mark variability and quality depending on the fundamental parameters: angle of attack, tool mark depth and substrate material [J]. Forensic Sci Int, 2015 (251): 40-49.

[8] Baiker M, Petraco N, Gambino C, et al. Virtual and simulated striated tool marks for forensic applications [J]. Forensic Sci Int 2016 (261): 43-52.

[9] Nicholas D, Petraco P, Shenkin J, et al. Addressing the national academy of sciences' challenge: a method for statistical pattern comparison of striated tool marks [J]. J Forensic Sci, 2012, 57 (4): 900-911.

[10] Bachrach B, Jain A, Jung S, et al. A statistical validation of the individuality and repeatability of striated tool marks: screwdrivers and tongue and groove pliers [J]. J Forensic Sci, 2010, 55 (2): 348-357.

[11] Taroni F, Champod C, Margot P. Statistics: A Future in Tool Marks Comparison [J]. AFTE J, 1996, 28 (4): 222-232.

于经验贝叶斯统计思想用于评价工具痕迹比对结果，认为贝叶斯统计理论用于工具痕迹检验是一种有希望的思想和方法。Petraco 等[1]使用共焦距显微镜采集弹壳痕迹、螺丝刀和凿子制作的线条痕迹的图像数据，使用主成分分析、规范变量分析和支持向量机模式识别方法检验工具痕迹，并用经验贝叶斯公式预测检验结果位于［0%，3%］区间的错误鉴定率的置信度为 95%。Keglevic 等[2]使用卷积神经网络分析螺丝刀线条痕迹，该方法在工具与客体接触角于 15–45 度的范围内痕迹的识别率可达到 95% 的平均精度。Spotts 等[3]对 50 个连续制造的鲤鱼钳制作的剪切痕迹，使用优化—证实算法分析了痕迹的相关性，验证了剪切痕迹具有唯一性。

线条痕迹的计算机定量化检验研究中，大多通过共聚焦显微镜或三维扫描设备采集痕迹的 3D 形貌数据，对痕迹的 3D 数据进行分析。相对 2D 数据，3D 数据能够清晰反映线条的高度、宽度和深度信息，且不受光照角度和强度的影响。但是，3D 数据采集也有其不足之处，如 3D 采集设备比较昂贵，且因其高分辨率而造成对较大面积痕迹的采集效率变低，也会产生较多的痕迹冗余信息。剪切痕迹整体形态近似半圆形，痕迹中具有较多非连续性线条，一般在痕迹检验分析时，需要人工预先设定剪切痕迹的短轴和长轴的位置。人工预先设定参数会使检验过程变得复杂，对检验结果也会造成一定的不确定性。本文针对痕迹的 2D 数据，使用对光照条件具有较好鲁棒性的局部二值模式（LBP）直方图及其衍生算子提取痕迹特征，可以克服痕迹 2D 图像采集时光源照射角度和强度不稳定的缺点，使用平衡误差好、能够处理大量输入变量和分类效果好的随机森林模式识别算法对痕迹进行分类识别，可以达到较高的识别率，对剪切痕迹的检验也不需要预先设定剪切痕迹的位置参数，使得检验过程简单方便。

二、方法

（一）实验痕迹制作与数据集采集

选择 10 个没使用过的螺丝刀（刀口宽 6 mm）、钢丝钳（刃口长 25 mm）和 8 个断线钳（刃口长 30 mm）作为实验工具。铅是一种延展性好、硬度低的金属材料，能够较好地复制工具作用部位上的特征，选用铅片和铅丝（直径为 5 mm）作为实验客体材料。

钢丝钳和断线钳由 2 瓣对称钢质的剪切刃组装而成，当实施其剪切功能时，能够在客体剪切端面形成剪切痕迹。钢丝钳和断线钳由于其具有强大的剪切能力以及携带方便的特点，在犯罪现场常被用于剪切锁梁、门窗防盗网、电线电缆等障碍物和作案目标，在被剪切物断面形成剪切痕迹。剪切工具实施一次剪切，会形成两个断头和四个剪切坡面。工具的剪切刃的每个剪切面在剪切客体时分别对应形成一个痕迹剪切坡面。剪切刃的每个剪切面相当于一个独立工具，即一个工具形成的四个痕迹剪切坡面相当于是由四个不同工具形成，且每个坡面痕迹是不相同的。Spotts 在实验中也验证了这个结论。

在铅丝上制作剪切痕迹，痕迹的坡面近似半圆形。在剪切痕迹形成的过程中，工具的剪切刃面与痕迹坡面之间存在切削作用和挤压作用。切削作用在痕迹坡面的剪切力方向（工具刃口运动方向）上形成线条，钢丝钳刃口阶梯斜刃会形成不连续的线条；挤压作用会让刃侧面的加工纹理印压到痕迹表面上，如钢丝钳刃侧面的加工铣纹会印压到痕迹上形成印迹。因剪切痕迹形成机理与螺丝刀线条痕迹形成机理不同，剪切痕迹表面线条模式更为繁杂，给传统的人工比对检验带来了困难。影响螺丝刀线条痕迹的最主要因素包括螺丝刀与客体的接触角和客体材料性质。螺丝刀在铅片上制作线条痕迹，会形成一致性非常好的线条。钢丝钳和螺丝刀在铅丝和铅片上形成的痕迹如图 1 所示。

［1］Petraco D K，Chan H，Forest P R，et al. Application of machine learning to tool marks：statistically based methods for impression pattern comparisons［J］. National Institute of Justice，2012（5）：114–117.

［2］Keglevic M，Sablatnig R. Retrieval of striated toolmarks using convolutional neural networks［J］. IET Computer Vision，2017，11（7）：613–619.

［3］Spotts R，Chumbley L S，Ekstrand L，et al. Optimization of a statistical algorithm for objective comparison of toolmarks［J］. J Forensic Sci，2015，60（2）：303–314.

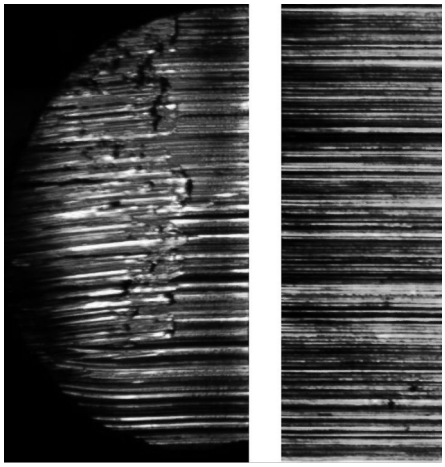


图 1　钢丝钳剪切痕迹（左）和螺丝刀线条痕迹（右）图例（钢丝钳剪切工具痕迹线条不连续，螺丝刀线条痕迹的线条是连续的）

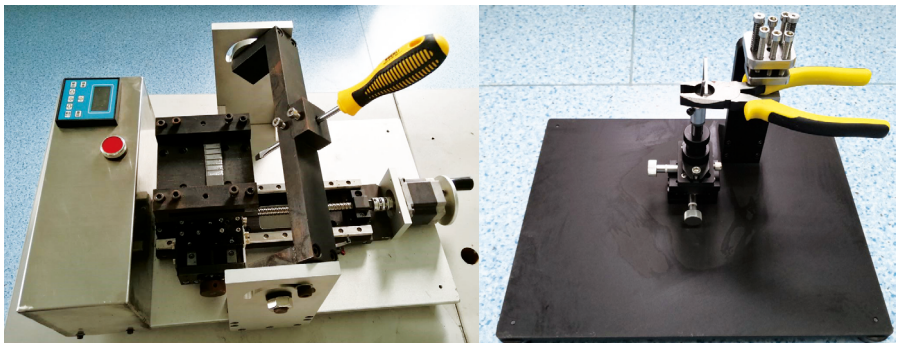


图 2　自制的螺丝刀线条痕迹制作装置（左）与剪切痕迹制作定位装置（右）

螺丝刀线条痕迹由自制的电动设备制作。螺丝刀和铅片固定在设备上，螺丝刀与材料之间接触角度以及刀口与材料之间的距离均可调节。实验中的接触角设定为 45°。每个螺丝刀制作 100 个痕迹，10 个螺丝刀共制作 1000 个痕迹。钢丝钳在制作剪切痕迹时，使用自制的剪切痕迹制作定位装置对刃口的剪切位置进行精度定位，每个钢丝钳制作 100 个痕迹，10 个钢丝钳共制作 1000 个痕迹，选择同一个刃侧面对应的剪切坡面作为实验痕迹。断线钳具有 2 级杠杆结构，剪切能力更强，经常被用于盗窃案中。断线钳在剪切过程中，其刃口在 2 级杠杆带动下产生位移，造成剪切刃定位误差，实验中刃口剪切位置定位精度误差约为 1–1.5 mm。每个断线钳制作 100 个痕迹，8 个断线钳共制作 800 个痕迹，选择同个刃侧面对应的剪切坡面作为实验痕迹。自制的螺丝刀线条痕迹制作装置与剪切痕迹制作定位装置如图 2 所示。

所有制作的实验痕迹由一位经验丰富的痕迹检验人员使用视频显微镜采集 2D 图像数据。采集图像时，显微镜放大倍数为 15 倍，图像分辨率为 1024 × 768（像素），光源的强度和照射角度自由调节但需保证采集的图像清晰。使用 Photoshop CS5 对采集的图像进行裁切。对于螺丝刀线条痕迹，靠近痕迹起始部位沿着痕迹整个宽度方向裁切图像，把裁切的图像保存大小为 250 × 720（像素）；对于剪切痕迹，只保留痕迹半圆形坡面区域，把裁切的图像保存大小为 250 × 720（像素）。裁切的图像均按照工具类型和类别进行标记，保存到实验数据集。

（二）机器学习线条痕迹检验

工具痕迹检验包含四个流程，如图 3 所示。首先构建实验痕迹图像数据集，然后对数据集中的数据进行滤除噪声预处理，实验中使用了高斯平滑滤除噪声，标准差取 σ =3 像素。

图 3　工具痕迹检验流程

工具痕迹的特征向量构建是痕迹识别比对的关键步骤，图像特征提取的效果直接影响识别比对结果的好坏。提取和构造痕迹数据的特征向量，包含三个步骤：四种 LBP 衍生算子特征提取、计算特征的直方图和构建数据的特征向量。LBP 由 Ojala 等[1]提出，对图像的每一个像素点，以该点的灰度值作为阈值，对其周围邻域 8 个像素进行二值化，并按照一定顺序将二值化结果组成一个二进制数，以此二进制数的值作为该点的响应。LBP 特征对单调的灰度变化具有不变性，但对灰度突变无能为力，这种以邻域像素值计算中心像素值的方式对椒盐噪声敏感，稳定性差。为了弥补此不足同时加快计算速度，Liao[2]提出了分块局部二进制模式（MBs_LBP）算子（其中 s 表示分块的大小）。在实验中，采用了 3×3 分块 LBP 算子，即 MB3_LBP 算子。

以 1 个像素的 8 邻域计算二进制数有 2^8=256 种模式，随着邻域内采样点数增加，直接采用 LBP 模式的统计量来表达图像信息时，数据量会过大，会造成痕迹特征维数剧增可能导致维数灾难的后果。对此问题，Ojala 提出一种“等价模式”（Uniform Pattern）的 LBP 方法来提升图像灰度的统计性，即 ULBP。

工具痕迹比对需要考虑痕迹局部特征的对应性，两个同一的工具痕迹其相同位置的局部特征应该相同。考虑到痕迹的局部特征与整体特征的关系，实验中把工具痕迹图像分成若干 8×8 的子区域。然后对各子区域进行统计得到等价化后灰度直方图，再将所有子区域的灰度直方图构建为一个特征向量。综合以上策略，实验分别使用了 LBP、ULBP、MB3_LBP 和 MB3_ULBP 四种算子对工具痕迹数据集进行特征提取并构建痕迹的特征向量，其流程如图 4 所示。

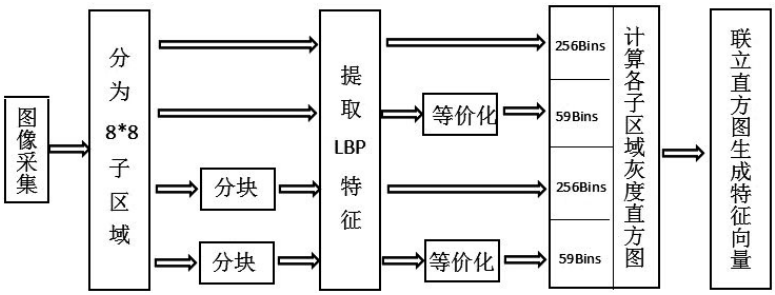


图 4　四种算子提取特征的流程图，由上至下依次为 LBP、ULBP、MB3_LBP 与 MB3_ULBP

最后，使用随机森林（Random Forest）分类器对所有痕迹的特征向量进监督训练。随机森林是一种基于多个决策树的集成学习方法，由 Breiman[3]提出。其与 Bagging[4]一样使用自主采样法生成多个训练样本子集，在每个训练样本子集上单独训练一个决策树，未被采样的训练样本则用来估计分类器的泛化能力。不同于 Bagging，随机森林训练中加入随机属性选择，对于基决策树的每个结点，先

[1] Ojala T，Pietikainen M，Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. Pattern Analysis & Machine Intelligence IEEE Transactions on，2002，24（7）：971-987.

[2] Liao S，Zhu X，Lei Z，et al. Learning multi-scale block local binary pattern for face recognition [C] //International Conference on Biometrics. Springer，Berlin，Heidelberg，2007：828-837.

[3] Breiman L. Random forests [J]. Machine Learning，2001（45）：5-32.

[4] Breiman L. Bagging predictors [J]. Machine Learning，1996（24）：123-140.

从该结点的属性集合中随机选择 k 个属性的子集，然后再从这个子集中选择一个最优属性用于划分。本文中随机参数取 $k=\log_2 d$（d 为特征向量维度）。在每个样本子集上生成对应的决策树后，这些独立同分布的决策树就构成了一个森林，对于一个新的输入样本，森林中的所有决策树被用来判断其归属分类，最终的类别取为这些决策树输出类别的众数。

三、实验与结果

分别使用 LBP、ULBP、MB3_LBP 和 MB3_ULBP 四种算子提取图像数据集特征并构建特征向量，然后将特征向量输入随机森林算法进行分类训练。所有的实验中，均对每个工具制作的痕迹数据集取其数量的 70% 作为训练数据，余下的 30% 作为交叉验证集。随机森林算法中，决策树样本和属性选择所涉及的随机种子取值均相同。图 5 所示四种算子处理螺丝刀线条痕迹的误差率随决策树数量的变化曲线，表 1 为四种算子处理螺丝刀线条痕迹的性能。

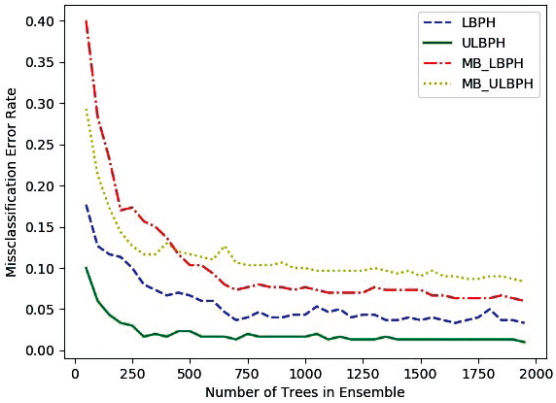


图 5　四种算子处理的螺丝刀痕迹的误分类率随决策树数量的变化曲线

表 1　四种算子处理螺丝刀线条痕迹的性能

	准确率（%）	特征向量维数	训练时间	单样本测试速度（S）
LBP	96.7	16384	长	0.92
ULBP	99.0	3776	短	1.11
MB3_LBP	94.0	16384	长	0.17
MB3_ULBP	91.7	3776	短	0.26

从对螺丝刀线条痕迹处理的检验准确率和综合性能来看，ULBP 算子效果最好。使用 ULBP 算子对三种工具制作痕迹的数据集做了四组实验，如表 2 所示。实验 1– 实验 3 分别单独对螺丝刀线条工具痕迹、钢丝钳剪切痕迹和断线钳剪切痕迹的数据集进行比对检验。在实验 4 中，把螺丝刀线条工具痕迹、钢丝钳剪切痕迹和断线钳剪切痕迹的数据集合成 1 个数据集进行比对检验。

表 2　四组实验

实验序号	实验数据集
1	螺丝刀线条痕迹
2	钢丝钳剪切痕迹
3	断线钳剪切痕迹
4	螺丝刀线条痕迹 + 钢丝钳剪切痕迹 + 断线钳剪切痕迹

（一）实验 1

本次实验对螺丝刀线条痕迹进行比对检验。痕迹的误分类率随决策树数量的变化曲线和痕迹比对结果的混淆矩阵如图 6 所示，痕迹的识别分类的准确率为 98.3%（准确率 = 正确分类的样本数 / 总样本数 × 100%；即：295/300 × 100%=98.3%）。混淆矩阵中，灰色小方框中的数字表示正确识别的痕迹数量，黑色小方框中的数字表示误分类的痕迹数量，坐标轴数字 1–10 表示为对应的 10 个螺丝刀制作的痕迹数据集。实验 2–3 的混淆矩阵中的数字表示均与此相似。

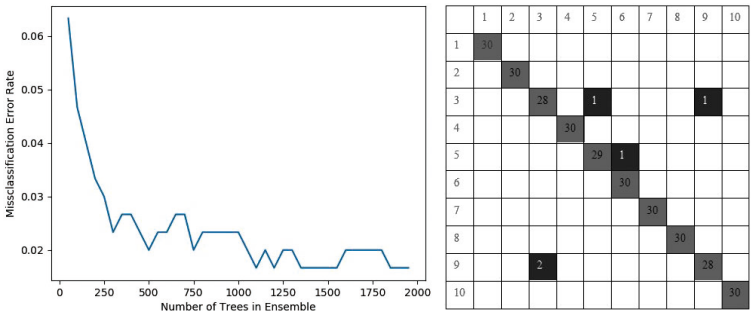


图 6　螺丝刀线条工具痕迹识别误差率随决策树数量的变化曲线（左）和痕迹识别分类的混淆矩阵

（二）实验 2

本次实验对钢丝钳剪切痕迹进行比对检验。痕迹的误分类率随决策树数量的变化曲线和痕迹分类结果的混淆矩阵如图 7 所示，痕迹的识别分类的准确率为 99.6%。

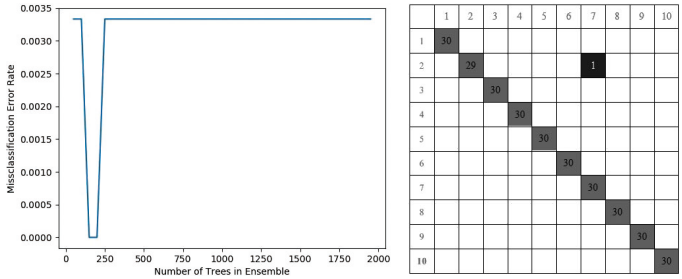


图 7　钢丝钳剪切工具痕迹识别误差率随决策树数量的变化曲线（左）和痕迹识别分类的混淆矩阵（右）

（三）实验 3

本次实验对断线钳剪切工具痕迹进行比对检验。痕迹的误分类率随决策树数量的变化曲线和痕迹的分类结果的混淆矩阵如图 8 所示，痕迹的识别分类的准确率为 81.3%。

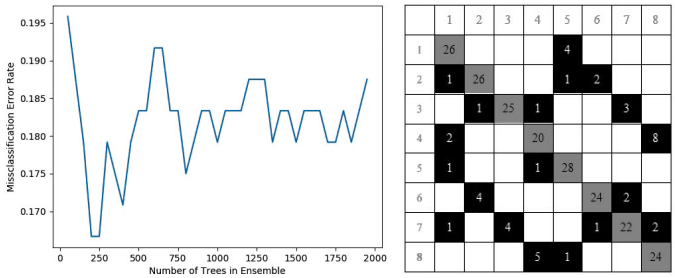


图 8　断线钳剪切工具痕迹识别误差率随决策树数量的变化曲线（左）和痕迹识别分类的混淆矩阵（右）

（四）实验 4

本次实验对螺丝刀线条痕迹、钢丝钳剪切痕迹和断线钳剪切痕迹数据集进行比对检验。痕迹的误分类率随决策树数量的变化曲线和痕迹的分类结果的混淆矩阵如图 9 所示，痕迹的识别分类的准确率为 93.7%。 混淆矩阵中，灰色小方框中的数字表示正确识别的痕迹数量，黑色小方框中的数字表示误分类的痕迹数量，坐标轴数字 1–10 表示为对应的 10 个螺丝刀制作的痕迹数据集，坐标轴数字 11–20 表示为对应的 10 个钢丝钳制作的痕迹数据集，坐标轴数字 21–28 表示为对应的 8 个断线钳制作的痕迹数据集。

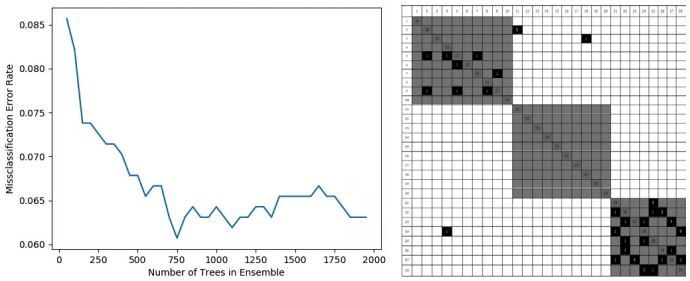


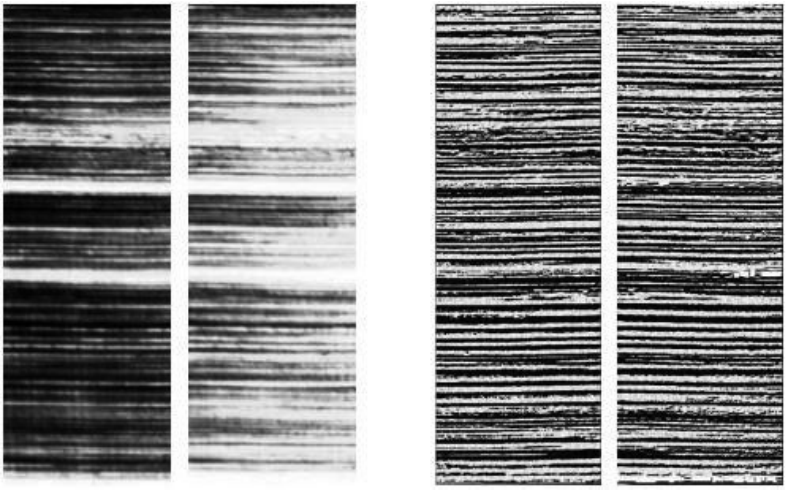
图 9　螺丝刀线条痕迹、钢丝钳剪切痕迹和断线钳剪切痕迹识别误差率随决策树数量的变化曲线（左）和痕迹识别分类的混淆矩阵（右）

四、讨论

从实验结果可看出，实验 1 和实验 2 分别对螺丝刀线条痕迹和钢丝钳剪切痕迹的识别比对准确率均在 98% 以上，比对结果优于痕迹的 3D 数据的识别比对结果。然而，实验 3 对断线钳剪切痕迹的识别比对的正确率只有 81.3%。实验 1 和实验 2 的结果好于实验 3，最主要的原因是，螺丝刀线条痕迹由自制的线条制作装置完成的，钢丝钳剪切痕迹由剪切痕迹定位制作装置完成，两者痕迹数据具有较好一致性。但是，断线钳的剪切刃由 2 级杠杆控制，在剪切时铅丝与剪切部位会产生相对位移，每次剪切位置会有误差，剪切痕迹的一致性较差，对识别比对造成影响。

实验 4 中对螺丝刀线条痕迹、钢丝钳剪切痕迹和断线钳剪切痕迹进行识别比对，综合识别比对正确率达到 93.7%。从实验 4 结果可看出，本方法对不同类型的工具痕迹综合数据的识别比对与单个类型的工具痕迹数据的识别比对的结果相差很小，绝大部分的错误分类仍在同一工具种类中，只有 4 个痕迹被错误的分类到其它工具种类中，如有 2 个 2 号螺丝刀线条痕迹误分为 11 号钢丝钳剪切痕迹，1 个 3 号螺丝刀线条痕迹误分为 18 号钢丝钳剪切痕迹，1 个 24 号断线钳剪切痕迹误分为 4 号螺丝刀线条痕迹。

本文实验中所有痕迹 2D 图像数据采集均没有对光照条件加以限制，即每个痕迹采集图像数据时光照强度和角度不相同。从实验结果可知，本方法对痕迹图像的识别比对具有较好的鲁棒性，对痕迹图像采集的光照条件具有较强的适应性，其主要原因是因为 LBP 算子具有光照不变性，图 10 所示同一痕迹在不同光照条件下采集的图像，经过 LBP 算子处理后的痕迹图像表现出很高的一致性。但是，采集图像时光照强度依然需要控制，光照强度过大或过小会造成图像的像素值为 255 或 0，从而造成痕迹信息丢失。



（a） （b）

图 10 （a）为同一痕迹在不同光照条件下采集的灰度图像，（b）为经过 LBP 算子处理后的图像

线条痕迹比对算法中，比较有代表性的研究如使用基于相关计算的 Wilcoxon 秩和检验的优化和证实两步法，使用人工定义的窗口在指定的痕迹区域计算 T1 值，通过 T1 值的大小判断两痕迹的相关性。该方法对工具痕迹的比对检验可达到较好效果。该方法中需要人工定义搜索窗口和痕迹的搜索位置，窗口的尺寸和痕迹的搜索位置会影响计算效率，甚至会给计算结果带来误差。对于线条一致性较好的螺丝刀线条痕迹，人工定义参数对计算结果影响可能不大。但是，对于具有非连续线条的剪切痕迹，需要人工定义剪切痕迹的短轴和长轴来确定痕迹的搜索位置，这样会给计算结果带来不确定性。显然，本研究的方法避免了人工参数的定义，使用痕迹图像纹理的统计特征和基于人工智能的模式识别方法，在痕迹的比对过程中减少了人为影响，比对过程简单。

五、结语

本文提出了一种基于人工智能机器学习的工具痕迹定量检验方法，主要包括工具痕迹数据采集与预处理、LBP 及其衍生算子特征提取，比对特征向量构造和随机森林识别分类。对 10 个没使用过螺丝刀制作的 1000 个线条痕迹、10 个钢丝钳制作的 1000 个剪切痕迹和 8 个断线钳制作的 800 个剪切痕迹的 2D 数据进行单独和综合比对检验实验。实验结果表明，该方法对痕迹图像采集的光照条件要求不高，方便痕迹的 2D 图像采集，对相同方式下形成的线条痕迹和类线条痕迹（剪切痕迹），均有较好的识别能力。本方法对工具痕迹定量化检验鉴定具有一定实用价值，下一步将继续研究制作痕迹定位误差与识别比对结果之间的规律问题。

（责任编辑：田　伟）

Study on Examination of Striated Toolmarks Based on Machine Learning

Yang Min[1,2]　Mou Li[1,2]　Fu Yiming[3]

1. Criminal Justice School in Zhongnan University of Economics and Law, Wuhan;
2. Forensic Sciences and Social Governance Disciplinary Innovation Base, Wuhan;
3. Department of Forensic Science in Guangdong Police College, Guangzhou

Abstract: The traditional method of manual examination of tool mark is challenged in the court for its subjectivity. With reference to the challenging, the computer-based approach have been studied in the world. The approach mainly focused on extraction of striation feature and statistical examination of striations. The machine learning method was studied, and four groups of experiments were conducted with a 2D image dataset of tool marks made by screwdrivers, cutting pliers and bolt clippers. The four LBP derivatives operators were developed to extract the tool-mark features and then construct the features into a feature vector. The random forest algorithm was adopted to identify the labeled feature vectors by supervised learning. The experimental results show that the proposed method achieved a high-rate of identification of the striated marks generated under identical conditions, and reduced the uncertainty of results examined by traditional method. Furthermore, the proposed method is immune to the unstable illumination when the image data of the striated marks are collected, and avoids the difficulty in mark inspection caused by manually preset parameters in the existing methods.

Key words: Forensic sciences; Striated mark; Artificial intelligence; Machine learning