arXiv:2303.04930v1 [cs.LG] 8 Mar 2023

基于核双样本检验的多模态多用户表面识别

Behnam Khojasteh，学生会员，*IEEE*，Friedrich Solowjow，Sebastian Trimpe，会员，*IEEE*，以及 Katherine J. Kuchenbecker，会士，*IEEE*

摘要—机器学习和深度学习已被广泛用于通过图像和时间序列接触数据对物理表面进行分类。然而，这些方法依赖人类专业知识，并涉及耗时的数据处理和参数调优过程。为克服这些挑战，我们提出了一种易于实现的框架，可直接处理异构数据源以完成分类任务。我们采用数据对数据的方法，通过对从多模态数据（例如图像、声音、触觉信号）中提取的两组数据进行核双样本检验，自动量化高维空间中分布的显著差异。鉴于视觉-音频-触觉表面识别这一应用具有工业相关性、难度较高且基线方法性能优异，我们通过与经过专家精心设计的分类器进行基准比较，展示了该技术的有效性；消融研究证实了处理流程中关键组件的实用价值。正如我们的开源代码所示，在包含 108 种表面类别的标准多用户数据集上，我们达到了 97.2% 的准确率；在更困难的任务版本中，性能比最先进的机器学习算法高出 6%。我们的分类器在标准算法设置下仅需极少的数据处理便能取得这一性能，这一事实进一步彰显了核方法在学习识别复杂模式方面的强大能力。

摘要—致从业者：我们演示如何将核双样本检验应用于表面识别任务，探讨改进机会，并说明如何将这一框架用于具有类似特性其他分类问题。自动化表面识别有助于表面检测和机器人操作。我们的算法量化类别相似性，因此会输出相似表面的排序列表。这项技术非常适合用于新收到的材料或新制造的零部件的质量保证和文档编制。更广泛地说，我们的自动化分类流程能够处理异构数据源，包括图像和

振动、力及其他物理信号的高频时间序列测量。由于我们的方法绕过了耗时的特征工程流程，专家和非专家都可以利用它实现高准确率分类。对于尚无现成模型和启发式方法的新问题，这种方法尤其具有吸引力。除了具备优良的理论性质外，该算法在实践中也易于使用，因为它只需要进行核函数评估。其透明的架构无需高昂的优化成本，即可针对不同传感组合，快速洞察给定用例。实践者还可以使用我们的流程来获得最短的数据采集时间

B. Khojasteh 和 K. J. Kuchenbecker 任职于德国斯图加特的马克斯·普朗克智能系统研究所（MPI-IS），同时也任职于德国斯图加特的斯图加特大学工程设计、生产工程与汽车工程学院。F. Solowjow 和 S. Trimpe 任职于德国亚琛的亚琛工业大学机械工程数据科学研究所，同时也任职于德国斯图加特的 MPI-IS。联系邮箱：khojasteh@is.mpg.de

本文已提交至 IEEE，拟供其考虑发表
版权可能会在不另行通知的情况下转让，此后此版本可能不再可访问。

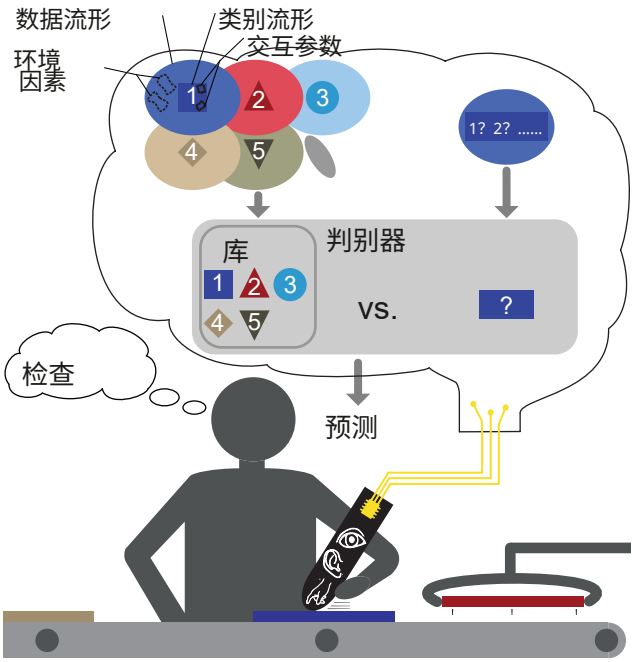


图 1. 我们针对表面识别任务提出的总体概念。由多模态视觉-听觉-触觉信息生成的数据流形包含环境因素（例如背景噪声）以及由交互参数（例如工具速度、法向力）张成的、针对不同表面的互补子空间。性能良好的判别器通过将未标记的检测试验与先前提取的类别流形库进行比较，提取出该试验的特征。

针对新传感器记录中的独立时间序列样本。

索引词—**自动化、分类、多模态数据、时间序列、核方法、双样本检验、触觉表面识别**

I. 引言

AUTOMATION 通过补充和提升人类工人的技能，增强工业系统的能力。自动化的一个新兴领域是表面识别：通过物理交互识别表面是一项颇具挑战的任务，与制造业和建筑业密切相关。人工感知系统可以收集并处理由人引导其与物理对象交互所产生的多模态数据，以帮助检查和识别其表面。为了模拟人类感知的丰富性，这些系统需要评估多样的表面属性 [1]。

表面探索产生的多模态传感器数据涵盖丰富的高维空间（图 1）。我们希望识别收集无标签多模态数据时所接触的表面，同时认识到所选交互参数和环境因素都会显著影响记录结果。对于视觉、音频和触觉模态，交互参数可包括照明、传感器位置、扫描速度、施加的法向力，以及传感工具的形状和属性；环境因素则可能包括温度或背景噪声。通用表面识别算法面临的主要挑战是检测多模态信息中固有的、与因果内容相关的属性，从而使这些特征能够从未见数据中检索出来。

人工系统如何有效利用多模态信息来推断表面属性，至今仍不明朗。近年来，表面识别领域的跨模态学习算法取得了进展，包括根据触觉 [2] 或触觉听觉线索 [3]推断图像，以及视觉到触觉感知 [4]。先前研究表明，融合多模态传感器信息 [5], [6]能够提升表面识别效果。然而，要确定哪种多模态数据组合最适合表面分类则更加困难，尤其是因为性能取决于所采用的算法。

如第 II 节所述，机器学习和深度学习均已用于表面分类。传统机器学习分类器依赖依赖问题的特征进行预测。确定合适的特征十分繁琐，需要专家知识来寻找与内容相关的特征，通过相关性分析去除冗余特征，检验特征质量以确保其在类内和类间具有唯一性，并在训练过程中反复优化特征。此类处理流程的理想结果是一组相互独立、具有描述性的特征，并且能够泛化到未见数据。然而，采用所选特征得到的最终结果在很大程度上取决于主观判断，因此难以评估给定方案是否会出现过拟合。深度学习算法无需特征工程，而是通过高效学习数据中的模式来表示所设定类别内的相关性。例如，仅使用图像训练的深度学习网络可以实现较高的表面分类准确率，但当应用于仅略微偏离训练分布的图像时，可能会产生严重的误分类 [7], [8]。多模态数据源可能有助于形成更稳健的表征 [6], [9]。除了普遍存在的过拟合问题，深度学习方法还需要大量训练数据，并且需要投入大量精力和专业知识来处理异构数据源及调整超参数。总之，自动化以往方法中耗时的流程，并尝试直接从多模态传感器数据中学习，有诸多充分理由。

我们通过利用 Gretton 等人的核双样本检验 [10]，实现了一种纯数据驱动的多模态分类方法。我们基于 Solowjow 等人开创性的工作，直接根据动力系统的时间序列输出对其进行分类，无需特征工程 [11]。本文的主要贡献如下：

- 我们改进并扩展了 Solowjow 等人的基于核的分类框架 [11]，使其适用于多模态多用户表面识别任务。我们提出的方法能够直接检查异构数据源，例如图像以及低频和高频时间序列。
- 我们使用 Strese 等人包含 108 种表面的多模态数据集 [12]，对算法在表面识别任务上的表现进行基准测试。尽管我们的分类设定更具挑战性、数据处理极少且领域知识有限，但该算法的表现优于已发表的、由专家精心设计的机器学习分类器结果。我们公开分享表面识别处理流程的代码，以便未来进行比较。
- 基于表面识别任务并结合消融研究，我们为如何将这一高效框架用于其他分类任务提供了指南。所提方法可直接应用于多种问题，包括异构数据源、时间自相关导致的虚假效应、训练数据稀缺、交互参数以及其他环境因素。

II. 相关工作

过去十年间，人们提出了许多用于物体和表面识别的算法。以下总结将表面分类研究分为两类常见方法：专家设计特征和深度学习特征。本节最后简要概述了我们为这项任务选择的工具——核双样本检验。

A. 使用机器学习分类器设计特征

关于物体和表面识别任务的人类受试者研究表明，人们会使用各种描述来对触觉感受进行分类。在触觉感知方面，研究人员已确定四个主要触觉维度（硬度、粗糙度、温暖度和摩擦力） [13], [14]。进一步研究将主要触觉维度细化为五个 [15]和十五个子维度 [16]。

Strese 等人 [6], [7], [17]对特征工程表面分类进行了广泛研究，并在 69（86%）、108（91.2%）和 186（90.2%）个表面上取得了令人印象深刻的识别准确率。对于 69 个和 108 个表面，分类使用一位专家的训练集和另外十个人的测试集；而对于 184 个表面，训练数据和测试数据均由同一人采集。众所周知，将分类推广到来自其他人的表面交互数据更具挑战性 [6], [17]–[19]。Strese 等人使用了传统机器学习分类器，例如 k近邻（k-NN）、随机森林，或多个分类器的组合。他们的特征来自听觉、视觉和触觉数据，以及超出人类感知能力的数据轨迹，例如红外（IR）光反射率和电导率。基于其最新的材料扫描系统，他们为多模态数据提出了 15 种感知特征 [7]。Burka 等人 [20] 的另一项研究使用相同的手工设计特征和多类支持向量机（SVM）分类器，将 69 表面数据集上的分类准确率提升至 87.5%。针对新的表面记录，Burka 等人还证明，其方法优于 Strese 等人的特征

在分类性能方面，相较于更简单的依赖扫描的特征，后者会随施加力和工具尖端速度而变化 [20]。Lu 等人 [19] 最近使用一种新颖的谱特征，在 69 表面数据集上对九种纹理类别的多用户听觉-触觉数据实现了 74% 的分类准确率；其朴素贝叶斯分类器使用了十位参与者的数据，并采用 7 比 3 的训练测试划分。

多项关于平面扫描机器人系统的研究显示，其表面分类准确率很高。例如，Sinapov 等人 [21] 使用 k近邻分类器和时频加速度特征；Jamali 等人 [22] 则提出了一种朴素贝叶斯分类器，使用时间均值以及从应变计读数中提取的显著频率特征。Fishel 和 Loeb [23] 根据仿生多流触觉传感器（SynTouch BioTac）进行的多次机器人控制扫描，计算统计特征，并使用朴素贝叶斯分类器，在 117 种纹理中实现了 95.4% 的识别准确率。

在机器人与三维物体交互领域，Chu 等人 [24] 表明，配备两个 BioTac 的机器人在使用 25 个触觉形容词为物体贴标签时，能够达到与人类平均水平相当的表现。他们实现了一种多核支持向量机分类器，提取标准静态特征和动态特征；后者由频谱振动的前四阶统计矩构成。Kaboli 和 Cheng [25] 将这一思路扩展到配备五个 BioTac 的机械手，并成功分类了 30 种手持物体（98%）和 120 种表面（100%）。

B. 使用深度学习分类器学习特征

学习表面类别的另一种方法是训练神经网络，使其直接在潜在空间中学习固有特征。这类神经网络通常包含输入层、输出层以及隐藏层，各层均可通过非线性数学运算相互作用。在存在非平稳纹理和测量条件变化的情况下，深度学习架构表现出强大的能力 [26]。

卷积神经网络已成功用于物体的基于图像的识别 [27]。一篇综述论文 [28] 总结指出，AlexNet、VGGM、VGGVD 和 TCNN 等先进网络结构能够准确分类现有规模最大的视觉纹理数据库（CUReT、KTH-TIPS2、ALOT）中 95% 至 98% 的数据 [29]。

超越纯视觉方法，Zheng 等人 [9] 将工具接触加速度与由一名专家采集的图像融合到一个全卷积网络中，该网络能够以 98% 的准确率对 69 种表面进行分类。具体而言，他们的深度学习分类器从加速度信号中提取时频信息，以便从触觉域中学习潜在特征。Joolee 等人 [30] 仅使用加速度处理了相同的分类设定；他们的多模型融合网络达到了 98.18% 的准确率，优于多种机器学习和深度学习方法。Gao 等人 [31]将混合卷积与循环网络应用于新的物体图像，以及 Chu 等人此前记录的、在对 53 个物体进行探索程序时采集的触觉数据 [24]；这种以学习为重点的方法优于基于人工设计特征的机器学习分类器。Strese 等人 [7] 其表现略优于（90.7% 对 90.2%）他们自己的机器学习分类器；该分类器融合了听觉、触觉和视觉模态的数据，而这里仅使用经过 AlexNet 改进卷积网络处理的图像 [27]。然而，由于缺乏有关接触属性的信息，他们性能更优的基于图像的深度神经网络在 184 种表面上出现了更多严重的误分类。错误的合理性是评判实际应用中所用分类技术的重要方式。Wei 等人 [32]最近仅使用听觉-触觉数据，在八种材料类别上实现了 87.6% 的分类准确率。他们的多模态卷积神经网络架构高效利用多尺度时间信息，对新的表面进行分类。

C. 使用核双样本检验进行学习 t

在高维空间中运行的核方法是机器学习中一种强大且广受欢迎的工具 [33]。除了像支持向量机那样映射单个数据点之外，还可以将完整分布直接映射到再生核希尔伯特空间（RKHS）中。由于其结构，希尔伯特空间配备了标量积，而标量积可通过核函数评估来计算。概率分布的嵌入称为核均值嵌入；Muandet 等人 [34]近期发表的综述全面总结了现有研究。具体而言，核均值嵌入为最大均值差异（MMD）提供了易于计算的近似，MMD 是概率分布空间上的一种度量。最大均值差异度量通过考虑低阶统计矩（例如均值、标准差、偏度、峰度）以及高阶统计矩，量化两个分布之间的差异。当不同类别之间的数据分布存在差异时，最大均值差异可作为分类任务的合适判别度量。Gretton 等人 [10] 提出了一种严谨的、基于最大均值差异的统计检验：其核心是估计两个数据集之间的最大均值差异，并检验它们是否来自同一分布这一原假设。他们针对数值示例和现实应用开发了这种所谓的核双样本检验。此后，核双样本检验凭借强大的理论保证和广泛的适用性，已成为概率机器学习中的一种重要方法 [34]–[38]。

核双样本检验已有多种扩展 [39]–[41]。特别是，Lu 等人 [42] 最近使用最大均值差异度量来量化真实表面信号与合成表面信号之间的相似性。此外，对于某些应用而言，要满足核双样本检验中数据独立同分布（i.i.d.）的假设可能尤为困难。对于表面识别任务，我们将处理非 i.i.d. 时间序列数据。因此，我们将借鉴先前工作 [11] 中的最新成果，该工作将核双样本检验扩展至具有非 i.i.d. 数据的动力系统。除了提供理论基础之外，Solowjow 等人还展示了其框架在一项使用低频周期性惯性测量单元信号、包含两个标签的步态分类任务中的效率。这种比较动力系统的新方法

与表面交互场景相关，尤其是因为最大均值差异可以考虑时间序列演化过程中的多个数据点。

III. 问题 formulation

我们的目标是对来自物理表面交互的、未见过的多模态传感器记录进行分类。用于此任务的算法需要处理异构且可能具有高维的数据（图 1）。具体而言，它需要从一组具有代表性的类别库中，检索无标签数据流形的表面特征。由于这些有意义的特征难以通过应用第一性原理或训练模型获得，因此需要采用更务实的方法。

在数学上，我们纯粹通过比较数据分布来处理表面识别任务。与 Solowjow 等人 [11] 的先前研究类似，我们将表面交互建模为平稳随机过程或动力系统的实现。我们假设一组 $C \in \mathbb{N}$ 独特的表面会分别产生不同的分布 $\mathbb{P}_1$、……、$\mathbb{P}_C$。我们的目标是将新的人类用户所采集的未标记表面记录，分配到相应的表面分布，以确定它们最有可能来自哪种表面类别 c。

IV. 所提方法

我们提出了一个识别框架（图 2(a)），以解决上述分类问题。我们假设每个表面类别 c 有 $Q \in \mathbb{N}$ 次试验，每次试验均包含来自 $S \in \mathbb{N}$ 个信息源的数据。因此，对于第 q 个表面试验 $Y^q = \{Y_1^q, \ldots, Y_S^q\}$，我们获得一组信息源。信息源可以包括来自不同传感器和/或由同一传感器通过不同探索程序获得的数据。在我们的分类设定中，表面试验 Y^q 和 Z^q 分别属于训练集和测试集。在我们的方法（图 2(b)）的核心部分，我们从两个具有未知分布 $\mathbb{P}_{Y_s^q}$ 和 $\mathbb{P}_{Z_s^q}$ 的数据流 Y_s^q 和 Z_s^q 中提取两个样本集 $[y_1, \ldots, y_n]$ 和 $[z_1, \ldots, z_m]$，其对应的样本量分别为 n 和 m。随后，我们量化给定信息源 s 中两个样本集的分布相似性。差异分数（DS），即 $\mathrm{DS}[Y^q, Z^q]$，将所有 S 个信息源计算得到的相似度统一起来。通过与类别库中所有数据进行成对比较，我们将试验 Z^q 分配给某个分布，进而分配给某个表面类别 c。

对于表面识别任务，我们考虑一个公开的工具——表面数据集 [12] 由十一名受试者记录的图像和时间序列数据组成。我们的识别流程（图 2）仅依赖一名专家测量数据构成的离线数据库；在这一被动训练阶段无需进行特征提取。在测试阶段，我们将数据库中每种表面的数据分布与另外十名用户的无标签测量数据进行比较，即一名专家对十名用户。我们的框架能够适配异构输入源、处理时间自相关，并补偿用户和会话相关效应，从而做出良好的分类决策。以下小节将详细介绍我们基于核双样本检验的识别范式的功能。附录 A 提供了数据集和代码的链接。

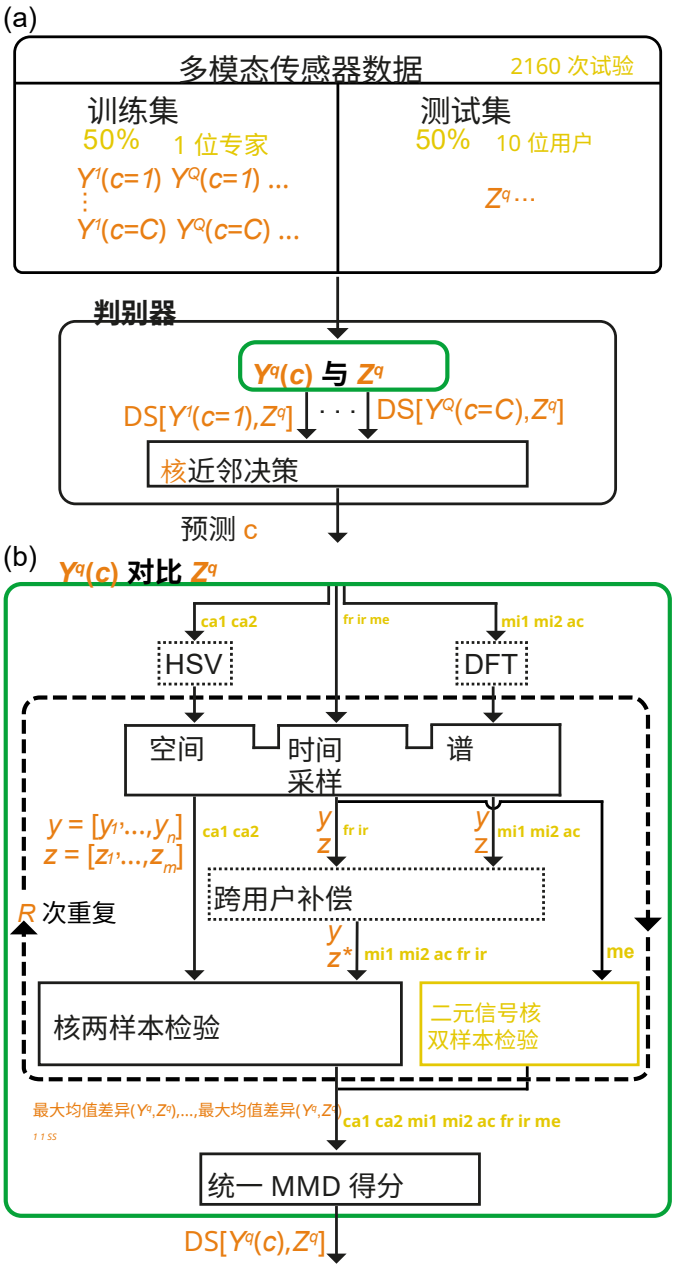


图 2. 所提出的表面识别流水线可直接处理数据，无需特征工程。橙色突出显示引入的术语，黄色突出显示数据集特定信息。(a) 判别器处理多模态数据，以预测表面身份。训练集由一名用户的表面数据库构成，测试集则包含另外十名用户的数据。(b) 我们的方法的核心是判别器跨 S 信息源比较两个数据试验 $Y^q(c)$ 和 Z^q。为阐明识别流程的有效性，我们针对三个虚线框所示的功能模块（HSV、离散傅里叶变换和跨用户补偿）进行了消融研究。

A. 概率分布空间上的度量

我们简要介绍本文所需的技术概念；详情请参阅 Muandet 等人撰写的综述论文 [34]。我们使用核均值嵌入来量化两个数据分布之间的差异。核均值嵌入将分布映射到更高维且可能是无限维的再生核希尔伯特空间中，在该空间中，某些任务变得易于处理。

在对核 k 作出一些温和假设的条件下，可以证明该嵌入是单射，从而确保只有一个对象会被映射到零；这一性质是进行有意义的统计检验的基础 [34]。这类核称为特征核。设 $\mathcal{H}$ 为一个再生核希尔伯特空间， $\mathcal{Y}$ 为数据输入空间， $\phi: \mathcal{Y} \to \mathcal{H}$ 为数据点的典范特征映射，y_i, $z_i \in \mathcal{Y}$，使得 $\phi(y)$ $\phi(z) \in \mathcal{H}$。由于通常不知道映射 ϕ 到再生核希尔伯特空间的具体形式，所谓的核技巧便可用核函数 k 的计算结果替代内积，因此

$$\langle \phi(y), \phi(z) \rangle_{\mathcal{H}} =: k(y, z). \tag{1}$$

核双样本检验 [10]旨在根据来自这些分布的独立同分布样本，区分概率分布 $\mathbb{P}_Y$ 和 $\mathbb{P}_Z$。核双样本检验的核心是估计最大均值差异，其定义如下：

$$\begin{aligned} \mathrm{MMD}(\mathbb{P}_Y, \mathbb{P}_Z, \mathcal{H}) &= \sup_{\|f\|_{\mathcal{H}} \leq 1} \mathbb{E}_{\mathbb{P}_Y}[f(y)] - \mathbb{E}_{\mathbb{P}_Z}[f(z)] \\ &= \|\mu_{\mathbb{P}_Y} - \mu_{\mathbb{P}_Z}\|_{\mathcal{H}} \end{aligned} \tag{2}$$

它是概率分布空间上的一种度量。为了比较分布，最大均值差异考虑再生核希尔伯特空间 $\mathcal{H}$ （由核 k诱导）的单位球中的函数 f。等价地，我们可以用核均值嵌入 $\mu_{\mathbb{P}_Y}$ 和 $\mu_{\mathbb{P}_Z}$来表示最大均值差异，从而得到第二个等式（详情见 [34] ）。核均值嵌入是一种强大的统计度量，因为它包含了分布的所有统计矩。

通常，直接计算式（2）中的最大均值差异是不可行的。不过，有一些基于核的估计量可以高效地近似最大均值差异。此外，还可以根据集中性结果量化收敛速度。因此，可以构造高置信度置信区间和显著性水平为 α 的统计检验。我们使用 Gretton 等人 [10],提出的最大均值差异平方估计量。

$$\begin{aligned} \mathrm{MMD}_b^2[\mathbb{P}_Y, \mathbb{P}_Z] &= \frac{1}{n^2} \sum_{i,j=1}^{n} k(y_i, y_j) \\ &+ \frac{1}{m^2} \sum_{i,j=1}^{m} k(z_i, z_j) - \frac{2}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} k(y_i, z_j), \end{aligned} \tag{3}$$

其中， $[y_1, ..., y_n]$ 和 $[z_1, ..., z_m]$ 是独立同分布的随机变量。在我们的情形中，它们分别是来自表面数据流 Y_s^q和 Z_s^q的 n 和 m 样本，其分布 $\mathbb{P}_{Y_s^q}$和 $\mathbb{P}_{Z_s^q}$未知。双样本检验在可能的低阶和高阶统计量所构成的无限维特征空间中进行最大化，并自动选择差异最大的统计特性。无需预先了解 $\mathbb{P}_{Y_s^q}$和 $\mathbb{P}_{Z_s^q}$，也无需对其进行参数化，即可获得这两个分布之间的差异。

实际上，即使两个分布相同，由于抽样数据点数量有限，最大均值差异的估计值也会波动。双样本检验的阈值 κ 可确保在原假设（即两个分布没有差异）下，经验最大均值差异的置信界具有显著性水平 α 。如果式 (3) 中的经验平方最大均值差异满足

$$\mathrm{MMD}_b^2[\mathbb{P}_{Y_s^q}, \mathbb{P}_{Z_s^q}] > \kappa(n, \alpha), \tag{4}$$

则可以以很高的概率得出结论： $\mathbb{P}_Y \neq \mathbb{P}_Z$ $1 - \alpha$。

核函数 k 的选择至关重要；有许多有效的选择，其中可能包含可优化的超参数。本文采用一个常用的候选函数——平方指数函数，

$$k(y, z) = \exp\left(-\frac{\|y - z\|^2}{2\sigma^2}\right), \tag{5}$$

作为所有统计检验的核函数。对于长度尺度 $\sigma \in \mathbb{R}^+$的任意取值，我们都能得到一个有效的核函数。然而，由于我们考虑的数据量有限，算法的性能通常会受到超参数 σ取值的影响。我们不对σ进行优化，而是采用成熟的中位数启发式方法 [10]

$$\sigma^2 = 0.5 \cdot \mathrm{median}(\{\|z_i - z_j\|^2\}) \tag{6}$$

在一个样本子集（i、 $j = 1$、.....、100）中样本之间的距离。

B. 从时间序列数据中采样

由于公式（4）中的显著性检验要求数据相互独立，因此在使用时间序列数据进行核双样本检验之前，我们需要处理固有的自相关。时间序列中的自相关是序列依赖的一种类型，表示时间相关观测。在动力系统或随机过程中应对自相关颇具挑战，但这对于保留时间序列分析的理论保证至关重要。高阶模式识别任务尤其依赖于通过互信息等方法去除虚假自相关，正如先前关于皮肤摩擦模式的研究所示 [43]。

互信息和皮尔逊相关系数是衡量两个变量之间依赖关系的常用技术。在此，我们遵循先前研究中的与最大均值差异兼容的方法 [11]，并使用希尔伯特–施密特独立性准则（HSIC）来量化依赖关系。两个随机变量之间的 HSIC， Y 和 Z，可定义为 [44]

$$\mathrm{HSIC}(Y, Z) = \|\mathbb{P}_Y \otimes \mathbb{P}_Z - \mathbb{P}_{Y,Z}\|_{\mathrm{MMD}}, \tag{7}$$

在最大均值差异度量下，两个边缘分布$\mathbb{P}_Y$ 和 $\mathbb{P}_Z$的张量积 $\otimes$ 与它们的联合分布$\mathbb{P}_{Y,\ Z}$ 之间的最大均值差异。通常难以计算式（7）中定义的希尔伯特-施密特独立性准则。因此，我们使用核化估计 [45, 式（4）]

$$\mathrm{HSIC}_b(Y, Z) = \frac{1}{n^2} \mathrm{tr}(KHLH) \tag{8}$$

其中 H， K， $L \in \mathbb{R}^{n \times n}$ 以及核算子 $K_{i,\ j} = k(y_i,\ y_j)$ 和 $L_{i,\ j} = l(z_i,\ z_j)$。中心化矩阵

$$H = I - n^{-1} e e^\top \tag{9}$$

由单位矩阵 I 和全一向量 e得到。

在我们的设定中，时间序列测量数据表示与表面发生物理交互时的观测结果。我们假设，给定时间序列源 s 与表面 c 交互所产生的每次试验 $Y_s^q$with $q = 1, \ldots, Q$ 都是同一动力系统的独立过程实现。我们从每个 Q 个实现中提取两个间隔为 t 的数据点，即 $Y_s^q(t_1)$ 和 $Y_s^q(t_1 + t)$，并将其划分

算法 1 独立样本的最小距离

输入： $Y^{1,\dots,s}$ ^{Q}, R, t_1 $^{\text{最大值}}$, α

输出： 消失自相关时间 T^*

1: 初始化 $t = 0$ 和 $CB^+ > \kappa^-$
2: **当** $CB^+ > \kappa$ **时执行**
3: $t = t + 1$
4: **对于** $r \leftarrow 1$ 至 R，**执行**
5: 从 $[1, t_{1,\max}]$中随机初始化 t_1
6: 使用 t_1 和 t构建 Y 和 Z
7: 使用公式 (8) 计算希尔伯特-施密特独立性准则$_b(Y,\ Z)^r$
8: 计算 κ^r 基于 α 自助法的希尔伯特-施密特独立性准则$_b(Y,\ Z)^r$
9: **结束循环**
10: $\kappa \leftarrow$ 平均 $\kappa^{1,\dots,R}$
11: $CB^+ \leftarrow (1-\alpha)$ 分位数的 希尔伯特-施密特独立性准则$_b(Y,\ Z)^{1,\dots,R}$
12: **结束循环**
13: **返回** $T^* = t$

为两组数据 $Y := \{Y_s^1(t_1)\dots, Y_s^Q(t_1)\}$ 和 $Z := \{Y_s^1(t_1 + t)\dots, Y_s^Q(t_1 + t)\}$；按构造方式，这两组数据各自内部均为独立同分布。因此，式 (8) 中的 n 等于 Q。如果将相邻数据点之间的时间间隔 t 设得足够大，经验估计量希尔伯特-施密特独立性准则$_b$ 就会预测数据近似独立，这也与式 (4) 中的设定一致。

在实际应用中，计算近似独立的时间序列样本需要多个步骤；详见算法 1。在这一流程的核心部分，我们从所有以 Q 实现 t 间隔的数据点中提取 Y 和 Z，以计算希尔伯特-施密特独立性准则$_b$ 和阈值 κ。该阈值通过显著性水平为 α 的自助法方法得到。根据具体应用，可以修改参数 α，以控制统计检验期间的检验错误；例如，$\alpha = 0.05$ 将假阳性错误率限制在 5%。自助法过程（实现细节见附录 B）包括使用打乱后的矩阵 L，反复计算希尔伯特-施密特独立性准则$_b$，并根据 $(1-\alpha)$ 分位数推断出 κ。我们以随机起始时间 t_1 重复执行这一核心流程 R 次，以提高统计检验的准确率。通过以不断增大的时间间隔 t 执行上述步骤，我们最终将获得近似独立的数据。随后，我们的算法会输出一个最优时间间隔 T^*，确保对于所有 R 次重复，希尔伯特-施密特独立性准则检验功效 (CB^+) 的 95% 置信区间上界均低于平均检验阈值 κ。因此，在为最大均值差异检验提取时间序列数据的过程中，我们可以可靠地保证自相关逐渐消失。上界初始化范围 $t_{1,\ \max}$ 限制了开始提取子轨迹的位置；对于平稳过程，这一影响可以忽略不计。不过，参数 $t_{1,\ \max}$ 与时间序列总时长 T_{total} 一起，通过 $T \le T_{\text{total}} - t_{1,\ \max}$ 决定了最优 T^* 的搜索上界。

消失自相关时间 T^* 表示时间依赖性的混合速度，并揭示所观测动力系统的复杂性。因此，这些混合特性体现了时间序列背后过程的可预测性，类似于与熵相关的互信息。混合速度将在下文中进一步阐述，并应用于附录 B 中的表面信号样本。

C. 采样策略

核双样本检验受益于未经处理的数据通道，这有助于最大化差异；同时，众所周知，它在高维数据上计算效率很高 [10]。因此，我们主要将原始传感器读数输入检验，并避免滤波或降维等数据处理步骤（例如，多通道信号压缩）。

除了经典时间序列外，也可以将其他数据源输入核双样本检验。例如，图像中空间分布的像素值以及经傅里叶变换的数据具有很强的表达能力，数十年来一直应用于信号处理领域。特别是，频域数据具有随机独立性，因此天然适用于MMD检验。RGB（红、绿、蓝）和HSV（色相、饱和度、明度）等不同的颜色表示方式已广泛用于计算机视觉领域的图像分析；其中，HSV将纯色和亮度值等重要图像属性分别编码在不同通道中。

从概念上讲，可以将时间域和频域中的数据流 Y_s^q和 Z_s^q都输入核双样本检验，尽管这会带来更高的计算成本。然而，对于某些应用，数据在其中一个域中可能已具有足够的表达能力。我们提出根据数据源的均值是否编码了有意义的表面信息来选择采样域，也就是说，考察信号的直流分量。例如，图像的平均像素强度可能是表面身份的特征，而听觉和触觉振动等零均值数据源的高频波形则更具辨识度。我们依据这一标准提出等距采样和随机采样策略（图3），以提取用于MMD检验的数据点。

等距采样适用于信号均值非零的图像和时间序列，并以空间或时间方式进行（图 3(a) 和 (b)）。我们在初始图像区域或初始时间窗口内，随机选择位于 (a_0, b_0) 或 t_0 的第一个数据点。为简单起见，两个数据流中第一个样本的位置（y_1 和 z_1）可以相同。然后，我们分别以空间或时间距离 d_a 和 d_b 或 T 等距提取新的数据点，以尽可能广泛地覆盖数据。这种采样策略会为图像生成网格。对于时间序列，为避免自相关，两个连续数据点之间的距离应大于或等于 T^*。

对于具有高频特征的时间序列，我们在频域中采用随机采样方案（图 3(c)）。这种方法非常适用于在分类场景中具有类别特异性频谱、且频谱在数值上存在差异的数据。例如，工具与表面交互产生的加速度计读数呈现出纹理特异性频谱特征 [17], [46]。我们通过离散傅里叶变换（DFT）将信号流转换到频域。随后，我们在两个信号流相同且互不重复的频率位置上，随机提取一组绝对频谱幅值。由此，我们可以直接比较这些频率在两个数据集中的存在情况。

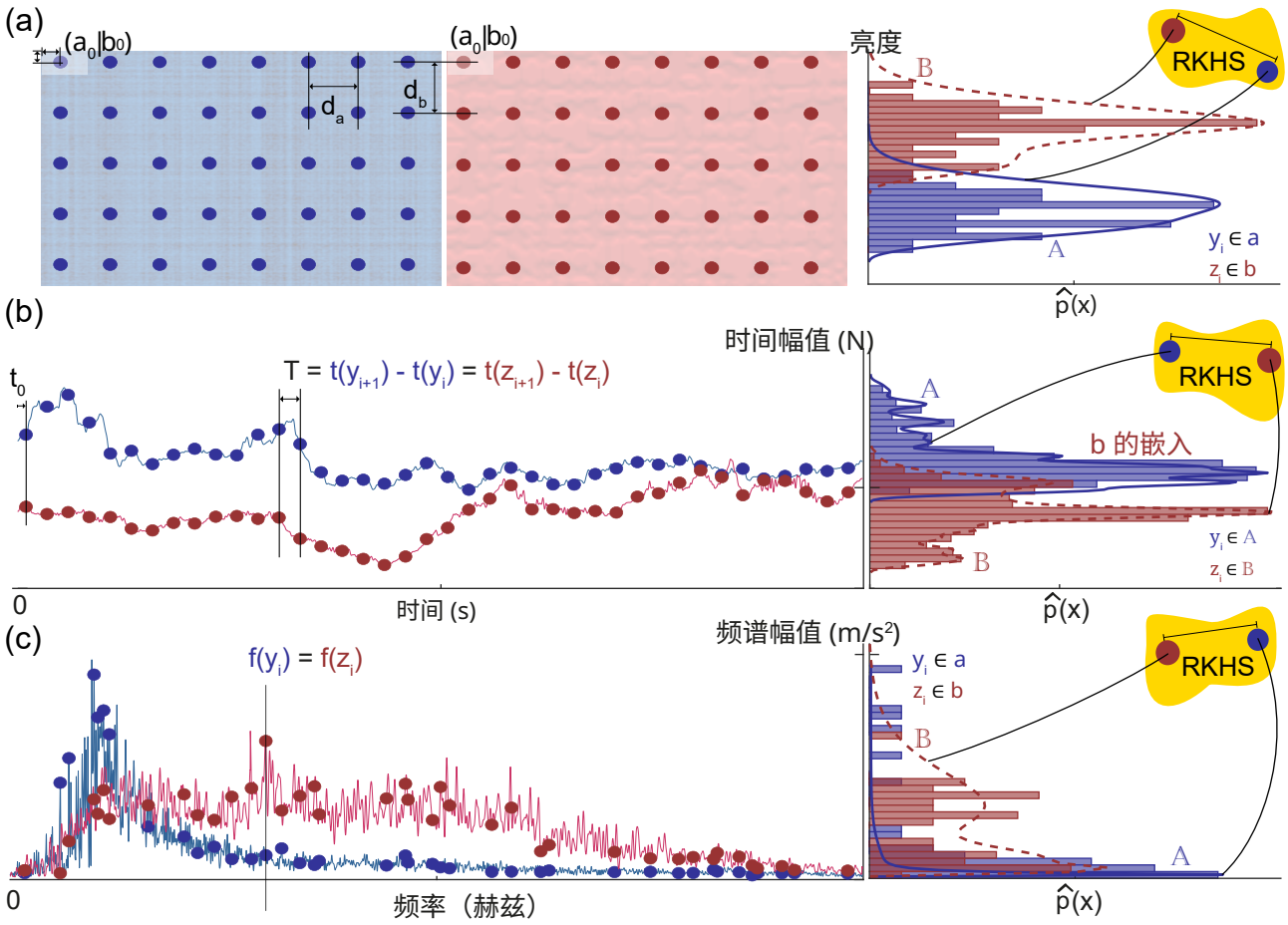


图 3. 我们用于核双样本检验的采样策略示意图，分别展示了 (a) 空间域、(b) 时间域和 (c) 频域。在图像和时间信号值上进行采样时，采用等间距方式；而在频域中，则采用均匀分布的随机唯一采样方案。我们在两个待比较信号的相同频率位置提取频谱幅值。核双样本检验将提取的样本点及其底层分布嵌入由内积构成的高维空间，即再生核希尔伯特空间（RKHS）。最大均值差异度量可自动确保最大化两个分布的可区分性。

核双样本检验的计算复杂度会随提取的数据点数量 n 和 m （式 (3)）呈非线性增长。为缩短计算时间，同时保持较高的分类准确率，我们使用较少的 n 和 m，反复执行 MMD 检验。在多次 MMD 检验中，输入数据点可能在初始采样和/或样本间距上有所不同。根据大数定律，各个最大均值差异估计的平均值会收敛到实际的最大均值差异。因此，我们能够更高效地处理数据，从而在我们的应用场景中获得更高的准确率。

D. 用户差异补偿

对于同一任务，不同人的传感器测量可能存在显著差异。例如，使用简单的纹理表面时，工具表面振动 [46] 和皮肤滑动摩擦力 [47]就会因人而异。这种用户间差异可能由扫描速度、施加的触摸力以及传感工具相对于表面的方向等因素造成。即使采用受控的测量协议，也很难消除接触力等信号中的用户间影响，以实现不受用户差异影响的模式分析，先进的信号处理技术 [47],和高阶非线性技术 [43]。

为了说明我们用于表面分类的方法能够推广到许多用户，我们需要减轻人的影响。我们采用一种整体性的统计方法来补偿个体差异；当数据集缺少用于建模用户差异的相关因果交互参数时，这种方法尤其有利。为了通过核双样本检验对不同用户的数据进行标准化，我们将从无标签数据流 Z_s^q中提取的数据点 $z = [z_1, \ldots, z_m]$的均值 $\bar{z}$

$$z^* = z + (\bar{y} - \bar{z}) \cdot e \tag{10}$$

移至库中数据流 Y_s^q的均值 $\bar{y}$ $[y_1, \ldots, y_n]$，并使用一个全一向量 e。这种跨用户补偿技巧背后的想法是，某些交互参数和环境因素的影响基本集中在分布均值中，而非高阶矩中；理想情况下，高阶矩中也包含有关表面属性的信息。

由于用户的动作及其他动力因素会显著影响基于接触的测量数据，我们对所有对这些用户动力效应敏感的信息源应用跨用户补偿，

例如，用户之间施加力或扫描速度等交互参数的变化，可能导致不同的摩擦力水平或振动谱能量。根据采样域的不同，该移位操作会对齐两个用户数据分布之间的时间信号均值或其谱能量。因此，在均值对齐的分布基础上，数据流比较仅关注高阶统计矩。

E. 多源分类

整合来自异构信息源的数据以作出集体分类决策并非易事，实现起来也可能十分繁琐。由于核的平移不变性以及我们对超参数 σ 采用的中位数启发式，MMD检验能够自动处理许多原本需要执行的预处理步骤。不过，不同维度和数据流中的数值大小各异，可能导致各信息源的MMD有所差异。因此，如何将各信息源的MMD 得分统一起来，形成整体分类决策，至关重要。

除了对同一比较的重复试验所得MMD求平均外，我们还分别计算每个信息源内的MMD，并将各信息源的数值合并为一个全局得分。根据式（3）中的MMD估计量，我们定义全局差异分数

$$\mathrm{DS}[Y^q, Z^q] = \prod_{s=1}^{S} \left(\mathrm{MMD}_b^2[\mathbb{P}_{Y_s^q}, \mathbb{P}_{Z_s^q}]\right)^{w_s} \tag{11}$$

用于比较两个试次 Y^q 和 Z^q。这一公式类似于加权几何平均数，或对各个对数变换值求算术平均数。该对数平均算子有助于实现不同信息源间MMD的尺度不变性。对于未标记表面的整体分类决策，每个信息源中的所有单独MMD 得分均可赋予权重 $w_s \in \mathbb{R}^+$。权重 w_s 使我们能够对某些信息源给予比其他信息源更高的重视；也可以为所有MMD 得分采用单位权重，正如我们在实验中所展示的那样。

*F. k*近邻分类决策

我们的识别框架中的判别器基于式（11）中的全局DS分数，依据 k近邻（k-NN）的概念进行分类预测。我们通过与库中某个表面具有最高概率一致性的方式，预测未标记表面的身份。根据库中包含 C 个表面的情况，未标记表面试验 Z^q 将被分类为类别 c，具体如下：

$$\min_{c \in C} \mathrm{DS}[Y^q(c), Z^q]; \tag{12}$$

它与全局 DS 最小的试验对相匹配，即最近邻（1-NN）。一般而言，也可以根据更多最近邻推断分类决策，这就需要在训练期间选择 k。

V. 实验

为展示我们的识别框架的效率，我们在 Strese 等人 [12] 提供的开放表面数据集上，将我们的算法与最先进的分类器进行基准测试。该数据集来自德国慕尼黑工业大学媒体技术教席（LMT）。我们证明，我们的统计方法无需学习表面类别，就能自动捕捉数据中最显著的表面属性，其表现优于采用专家设计特征的其他分类器。

A. 多模态表面数据集

LMT108数据集包含来自108种表面纹理的多模态测量数据（图4），这些数据由一支带有不锈钢工具尖端 [12] 的手持式传感器触针采集。108种表面各属于九种材料类别之一：（M）网格、（S）石材、（G）金属/玻璃/陶瓷、（W）木材、（R）橡胶、（C）地毯、（F）泡沫、（P）塑料/纸张，以及（T）纺织品/织物。每种类别包含的表面数量从五种到十七种不等。每次记录的表面试验包含九种信息源：两幅图像（有闪光灯和无闪光灯）；敲击动作产生的声音和加速度；以及工具拖动过程中产生的声音、加速度、摩擦力、红外表面反射率和金属检测数据。这些信息源来自六个传感器：相机（ca）、麦克风（mi）、加速度计（ac）、两个力敏电阻（fr）、红外表面反射率传感器（ir）和金属检测传感器（me）。除图像外，所有数据流都是时间的函数。对于每次试验，工具拖动产生的信息源涵盖4.8秒的数据，且没有突然的瞬态运动。加速度和图像各有三个通道（分别为坐标轴和颜色），力有两个通道。其余信息源均为一维。专家组和用户组都包含每种表面的十次试验（总计2160次），但采集数据的人数不同（一位专家与十位用户）。专家采用不同的动作、扫描速度和力度，以确保每种表面具有类内方差。其他十位用户则不受任何动作要求限制，自由移动触针。

我们注意到，无法考虑 LMT108数据集中的敲击加速度，因为十位用户均无法获得这些数据；在线档案包含专家的重复数据，而原始试次已丢失。因此，我们仅使用八个而非原来的九个信息源；样本数据见图 5。敲击产生的加速度瞬变包含有关表面的丰富且描述性的信息，可能有助于分类。尽管所考虑的信息源数量不一致，仍然可以进行定性比较。

B. 数据预处理

我们的识别框架（图 2）使用每个 LMT108 表面的八个信息源的所有传感器通道。我们用色相、饱和度和明度（HSV）颜色编码表示三通道图像。对于具有有意义直流分量的数据流，即



图 4. 在线 LMT108 数据集 [12]涵盖九类材料的 108 种表面：（M）13 种网格、（S）9 种石材、（G）9 种光泽材料、（W）13 种木质材料、（R）5 种橡胶、（C）15 种地毯、（F）12 种泡沫、（P）15 种塑料和纸张，以及（T）17 种纺织品和织物。

图像、力、红外反射率和金属扫描数据，我们采用等距采样方案。其余三个信息源（mi1、mi2、ac）表示具有不同交流波形的接触振动，因此我们在频域中进行随机采样。对于这些频谱听觉与触觉振动，我们在算法中采用与 LMT108 数据集的作者相同的频率范围（分别最高为 7500 赫兹和 1000 赫兹） [6], [12]。我们基于准则选择时间序列数据的采样域（时间域或频域），与原始研究中专家设计特征所使用的域一致 [6], [12]。此外，公式 (10) 中的跨用户补偿适用于所有受到以下一项或多项因素显著影响的信息源：扫描速度、施加力、工具动力学，或环境噪声等会话相关效应。对于 LMT108 数据集，我们对五个信息源应用跨用户技巧：mi1、mi2、ac、fr、ir。

C. MMD 测试设置

由于测量时长、采样率、数据维度以及时域或频域表示方式不同，八个信息源的数据流点数量各不相同。尽管数据规模不同，在每次单独的 MMD 计算过程中，我们都会从每个信息源的每个数据通道中采样 $n = m = 400$ 个数据点。提取的数据部分约占每幅图像的 0.3%、时间域力/反射率/金属信号的 0.8%、拖动声音的 0.9%、敲击声（44.1 kHz）的 14.4%，以及拖动加速度（10 kHz）的 6.1%。此外，我们针对每个信息源的所有表面试验比较，重复（$R = 10$）进行核双样本检验，并对 R MMD 得分取平均值。对于采用等距采样方案的五个信息源，从初始图像区域或时间窗口（针对随机选取的第一个数据点）中取得的数据点，不超过相应数据流点总数的 1.15%。

使用规定数量的数据点广泛覆盖数据

n 和 m 对于每项MMD检验，等间距采样策略使三个时间序列信息源（fr、ir、me）的采样间隔为 $T = 11.8$ 毫秒，两个图像的采样间隔分别为 $d_a = 17$ 像素和 $d_b = 18$像素（480×320 像素）。然而，采样时间间隔 T 并不满足所有力信号和红外反射表面信号的混合要求；金属信号大多只有一个信号值，因此不适合用于计算希尔伯特-施密特独立性准则。为使这两类信息源中的所有表面都满足自相关时间趋于零的理论保证，

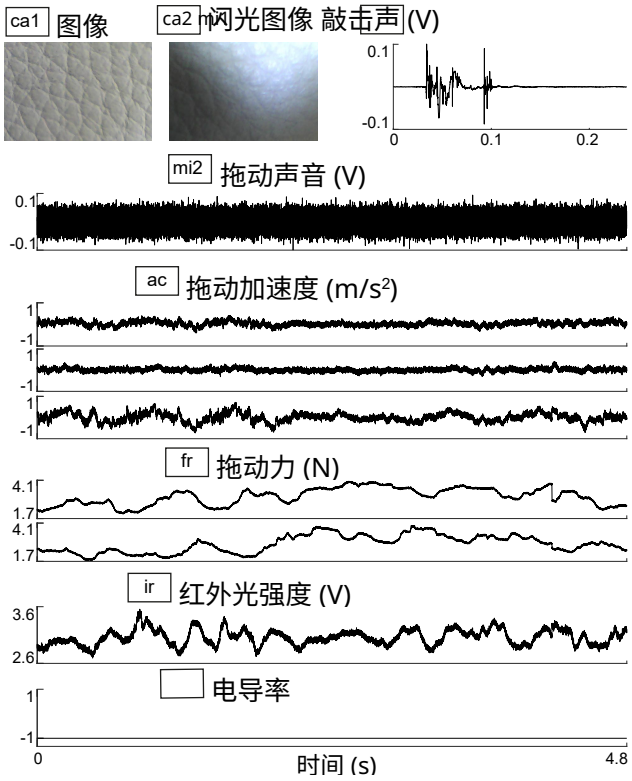


图 5. LMT108数据集 [12]中皮革表面的一次用户试验（测试阶段）的八个信息源。测量数据来自相机 (ca)、麦克风 (mi)、加速度计 (ac)、两个力敏电阻 (fr)、红外反射传感器 (ir) 和金属检测传感器 (me)。

要么需要大幅增加 T（减少采集的样本数），要么需要将测量时长延长至远超4.8秒。更长的时间序列测量有助于降低固有自相关，并且很可能提高分类准确率。尽管如此，为了平等对待所有信息源，我们仍采用每通道400个样本。尽管理论最优与实际选用的 T并不一致，实证结果将展示我们方法的有效性。

时间金属探测器读数为二进制值，即对表面中的导电材料进行瞬时检测。LMT108数据集包含十二种金属表面。通常，我们不会对二进制数据流使用MMD检验，而是仅关注低阶统计矩，例如此前作为分类特征考虑的信号均值 {v1}。

然而，为了以一致的算法架构平等处理所有信息源，我们也对金属检测读数应用核双样本检验。因此，当相同值数据会导致 $\sigma = 0$时，我们按照式（6）中的启发式方法设置核长度尺度 $\sigma = 1$。此外，我们将金属信息源的所有 MMD 得分均加上常数 1，以确保 $\mathrm{MMD}^{\mathrm{me}} \neq 0$ 在式（11）中使用几何平均数时不会造成信息损失。我们不对金属检测测量数据应用跨用户补偿，因为对齐分布均值会消除有关材料导电率的信息。

如前所述，式（11）中的加权几何平均数框架能够为不同信息源赋予不同的影响权重，并将其纳入单个最大均值差异得分，以作出整体分类决策。然而，此处我们不在训练过程中调整模型，因此简单地对所有信息源赋予相同权重，即 $w_s = w = 1$。

D. 消融研究

为展示我们提出的处理流程的功能，我们通过消融研究评估分类器的关键方面。我们分别移除识别流程（图 2）中的以下三个组成部分：图像的 RGB 到 HSV 转换、高频振动的 DFT 时域到频域转换，以及对五个选定时间序列信息源进行的跨用户补偿。因此，在 HSV 消融研究中使用标准 RGB 图像；在 DFT 消融研究中，振动信号（mi1、mi2、ac）采用等间隔时间采样，而非频谱采样。我们的可复现代码确保在启用或移除某项功能时采样相同的数据点，从而能够观察该功能本身的影响。为深入分析，我们评估了包含完整多模态信息以及仅包含单个信息源时，在启用或移除某项功能情况下的分类器性能。此外，我们还进行了另一种训练测试划分：仅基于专家试验进行表面识别（1 名专家对 1 名专家），以主要阐明跨用户补偿的有效性。为进一步深入分析，我们也报告了仅专家试验分类场景下另外两项消融研究的性能。由于仅有一个人的信号方差较小，1 对 1 的数据划分使分类任务更简单，因此使用多模态信息时能够获得（近乎）完美的识别率。因此，我们仅报告单个信息源下的分类器性能。

E. 对比方法

下面简要介绍此前发表的六种分类器，我们将用它们与自己的方法进行比较。这些分类器都使用专家精心设计的特征，并广泛用于表面分类；详情见 [48], [49]。朴素贝叶斯分类器是一类概率分类器，假设特征之间具有很强的独立性。如果独立性假设成立，与其他机器学习模型相比，该算法只需较少的训练数据便能取得良好效果。该分类器会选择给定观测与类别出现情况的联合似然概率最高的类别。高斯朴素贝叶斯模型进一步假设联合似然服从正态分布。

支持向量机分类器是一种判别式模型，它从训练集学习决策函数，将输入分配到各个类别。训练阶段可以包括映射到更高维空间，例如通过式（1）中的核技巧，最大化两个类别之间的间隔，并适应非线性决策边界。多类支持向量机使用多个二元分类来进行整体类别预测。与我们的方法相比，主要区别在于，映射到再生核希尔伯特空间的是单个数据点，而非整个分布。

给定测试集中的一个数据点，非参数 k-NN 算法会考察该点与训练集中 k个最近邻之间的特征距离，并通过相对多数投票得出预测结果。 k-NN 分类器在所用度量（例如欧几里得距离或马氏距离）以及最近邻的加权方案方面可能有所不同。经典的 k-NN 和我们的方法都基于训练集与测试集之间的最近邻概念进行预测；然而，与我们直接根据输入数据量化分布差异的方法不同，k-NN 考察的是显式特征之间的距离。

集成分类器采用决策树的基本概念，并结合多个分类器进行整体预测，即采用不同的分类器，或采用配置不同的同一分类器。这类算法会从特征空间中选取一组特征和实例，并对实例进行复制或加权，以供各个分类器进行最终多数投票。这种在训练前改变输入特征空间的操作也称为自助聚合，能够增强分类器的稳健性。

随机森林属于集成方法，使用自助采样特征空间。在此过程中，会随机选取一组较小的特征来训练每棵决策树。对于一个测试实例，所有树的预测结果通过多数投票机制决定最终结果。由于这类分类器需要严格训练，因此在大型数据集上计算开销很大。

另一种流行的集成分类器是自适应提升（AdaBoost）；它使用许多仅基于单个特征进行划分的决策树。在训练过程中，每棵新树都会考虑前一棵树的错误，以减少剩余的错误分类观测值。与随机森林中树的权重相同不同，错误较少的 AdaBoost 树会获得更高的最终投票权重，因此多个弱分类器可以组成效果更好的整体分类器。

F. 性能指标

Strese 等人 [6], [12] 在 LMT108数据集上使用上述基准分类器进行了 10折交叉验证（每折：1 位专家用户 + 9对 1 位用户）。相比之下，我们的识别流程在训练阶段仅将专家试次（占 LMT108 数据的一半）用于构建库，因此所处理的分类设定要困难得多（1 位专家对 10 位用户）。Strese 等人最新发表的论文 [6]通过以下方式评估了 LMT108 分类器的性能：

10 折分类准确率的均值和标准差。在 Strese 等人之前发表的论文 [12] 中，他们还通过分类精确率评估了一种基于欧几里得距离的 k-NN 分类器。为便于直接比较，我们对 LMT108 表面识别任务采用相同的两项性能指标。

类别 c 的分类准确率为

$$\mathrm{Accuracy}(c) = \frac{\mathrm{TP}(c) + \mathrm{TN}(c)}{\mathrm{TP}(c) + \mathrm{TN}(c) + \mathrm{FP}(c) + \mathrm{FN}(c)}, \quad (13)$$

其中，真正例和真负例分别为 TP(c) 和 TN(c)，假正例和假负例分别为 FP(c) 和 FN(c)。我们计算十个用户特定折的分类准确率（每个测试集包含 108 个试次），并求其均值和标准差，以得到总体准确率。此外，表面类别 c 的分类精确率为

$$\mathrm{Precision}(c) = \frac{\mathrm{TP}(c)}{\mathrm{TP}(c) + \mathrm{FP}(c)}. \quad (14)$$

类别精确率表示类别 c 的所有预测中，正确预测所占的比例。总体精确率通过对所有类别特定值取平均得到。

G. 结果与讨论

如表 I 所示，使用完整分类器，我们对 108 种表面类别的准确率达到 97.2%，尽管少使用了一个信息源，仍比次优模型（随机森林）高出 6%。对于全部 108 种表面，我们的平均分类精确率也达到 98%，比基于欧几里得距离的 k-NN 分类器 [12] 高出 12%。因此，我们的分类器具有很高的阳性预测能力。这两项关键结果表明，与所有先前开发的、使用基于专家知识的特征的分类器相比，采用我们通过检测最大分布差异实现的自动化方法进行表面分类效率更高。

表 I 使用 LMT108 数据集对专家设计的分类器与采用 S 个信息源的不同方法变体进行比较：准确率（均值 ± 标准差）和精确率

算法	准确率 [%]	S	传感器
高斯朴素贝叶斯 [6]	78.1 ± 0.3	9	6
支持向量机（径向基函数） [6]	79.1 ± 0.5	9	6
k近邻（马氏） [6]	84.1 ± 0.4	9	6
AdaBoost [6]	86.8 ± 0.8	9	6
随机森林 [6]	91.2 ± 0.8	9	6
我们的：视觉（ca1+ca2）	74.8 ± 2.5	2	1
我们的：听觉（mi1+mi2）	54.8 ± 16.3	2	1
我们的：触觉（ac+fr）	61.6 ± 6.0	2	2
我们的方法：视觉+听觉 (ca1+ca2+mi1+mi2)	84.1 ± 7.3	4	2
我们的方法：视觉+触觉 (ca1+ca2+ac+fr)	92.7 ± 2.2	4	3
我们的方法：听觉+触觉 (mi1+mi2+ac+fr)	74.2 ± 10.9	4	3
我们的方法：全部（ca1+ca2+mi1+mi2+ac+fr+ir+me）	**97.2 ± 2.1**	**8**	**6**
我们的方法：全部 – HSV颜色消融	95.8 ± 2.7	8	6
我们的方法：全部 – 离散傅里叶变换消融	95.1 ± 3.3	8	6
我们的方法：全部 – 跨用户补偿消融	88.7 ± 6.7	8	6
	精确率 [%]		
k近邻（欧几里得） [12]	86	9	6
我们的方法：全部（ca1+ca2+mi1+mi2+ac+fr+ir+me）	**98**	**8**	**6**

利用图像高效进行表面分类：关于表 I 所示的单模态表面分类性能，视觉模态（两幅三通道图像，ca1+ca2）效率很高。只用两幅图像，我们的分类准确率就几乎达到使用全部数据流的多模态高斯朴素贝叶斯分类器的水平。在颜色消融研究中，将 HSV 图像换成 RGB颜色空间图像后，识别准确率略低（95.8% ±2.7%），但与完整分类器相比，我们的方法受此变化的影响很小。然而，在两种训练-测试划分下，无论是否进行跨用户补偿，HSV 图像在单图像分类中的表现都显著优于其 RGB 表示（表 II）。我们认为，HSV 对单模态分类更为重要，因为这种色彩空间将有意义的表面信息封装在彼此独立的数据通道中。HSV 能降低对图像照明变化的敏感度，从而实现更稳健的图像分类。我们还观察到，在两种色彩空间、两种训练-测试划分下，无论是否进行用户补偿，标准HSV图像（ca1）的表现都优于闪光图像（ca2）（表 II）。相机闪光灯虽能揭示有关表面反射率的有用信息，但未照明表面图像的一些描述性属性似乎会因此丢失，不利于识别任务。此外，使用常规 HSV 图像（ca1）进行单通道识别的结果表明，图像色调（42%）和饱和度（34%）比亮度值（27%）包含更多信息。而在闪光图像中，这种差异不再显著（H：19%，S：18%，V：17%），因为光照条件使三个图像通道中的描述性信息趋于均衡。

时域与频域时间序列采样：离散傅里叶变换消融研究表明，使用时域采样处理高频振动源时，我们的完整分类器表现仅略差（95.1% ±3.3%）（表 I）。针对单个信息源的消融研究表明，在两种训练-测试划分和两种用户补偿设置下，对于大多数信息源，我们首选的基于准则的采样域表现往往更好（表 II）。如果对同一训练数据在时域和频域的数据上都执行最大均值差异检验，识别准确率会略高，达到 98.2% ±1.7%。接下来，我们将时间序列信息源从六个增加到十二个，并对时域和频域中的相同五个信息源采用跨用户技巧。由于这一性能提升需要对时间序列进行两倍数量的核函数评估，因此我们得出结论：采用单一基于准则的采样域已经非常高效。

跨用户波动来源：关于用户对分类准确率的影响，与此前发表的机器学习分类器相比，我们的算法中用户特定平均准确率的变异性更大。在我们的最优设置（$S = 8$）下，每位用户的测试集准确率介于 92% 至 99% 之间（七位用户为 $\geq 97\%$）。这种较大的差异源于我们在交叉验证中选择仅将专家数据纳入库，而未加入其他九位用户的额外数据。这些波动来自依赖会话的影响（例如环境噪声）和依赖用户的影响（例如动作和敲击力度）

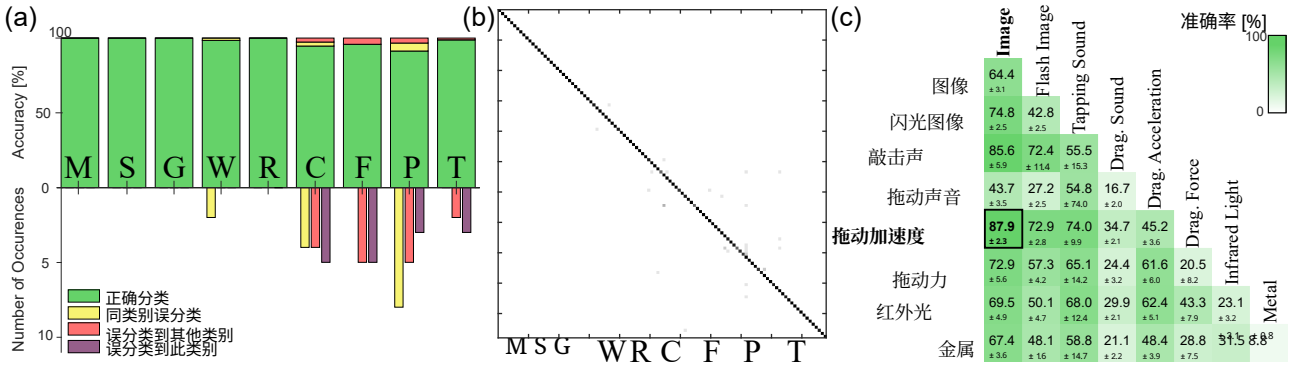


图 6。我们的方法的预测结果：(a) 分类准确率的相对统计值和绝对统计值；(b) 使用全部八个信息源（$S = 8$）对 108 个表面类别（$C = 108$）进行分类所得的相应混淆矩阵。为更好地展示相似纹理类型，我们按照先前使用的材料分类体系，对 108 个表面的结果进行了分组。(c) 单个信息源及其两两组合的分类准确率。表面图像和接触诱发的拖动加速度（框内所示）为表面识别提供了互补且稳健的信息源。

这些影响使表面数据更加多样，因此，使用较少人类受试者进行训练后，对未见数据的预测会更困难。触觉和听觉模态表现出特别大的用户差异。听觉信息源的差异大于触觉信息源，这可能是因为音频信号

表 II 两种训练-测试划分下单个信息源消融研究的分类准确率

信息来源	训练分割	跨用户补偿	准确率 [%] 时间采样	准确率 [%] 谱采样
mi1	1 位专家 vs. 10 位用户	no	50.9 ± 16.4	48.4 ± 18.3
		yes	47.6 ± 18.0	*55.9 ± 15.4*
	1 位专家对 1 位专家	no	22.4 ± 5.0	19.0 ± 3.1
		yes	21.1 ± 3.5	27.1 ± 3.7
mi2	1 位专家 vs. 10 位用户	no	4.2 ± 1.2	6.6 ± 1.3
		yes	4.3 ± 1.5	*15.3 ± 3.3*
	1 位专家对 1 位专家	no	12.6 ± 2.9	39.9 ± 4.3
		yes	13.0 ± 2.6	41.0 ± 2.7
ac	1 位专家 vs. 10 位用户	no	48.3 ± 5.6	26.4 ± 1.8
		yes	45.8 ± 4.5	*43.8 ± 2.3*
	1 位专家对比 1 位专家	no	46.9 ± 9.8	72.8 ± 6.9
		yes	52.6 ± 7.5	56.6 ± 7.5
fr	1 位专家对比 10 位用户	no	14.2 ± 3.9	5.6 ± 1.7
		yes	*20.2 ± 7.5*	7.1 ± 2.2
	1 位专家 vs. 1 位专家	no	33.5 ± 12.2	22.0 ± 6.3
		yes	12.4 ± 4.8	10.6 ± 4.1
ir	1 位专家对比 10 位用户	no	23.7 ± 4.6	8.0 ± 2.1
		yes	*23.5 ± 3.7*	12.3 ± 2.8
	1 位专家对 1 位专家	no	55.5 ± 5.1	28.0 ± 4.3
		yes	23.8 ± 4.5	24.9 ± 2.7
me	1 位专家 vs. 10 位用户	no	*8.5 ± 1.3*	8.1 ± 1.2
		yes	7.6 ± 1.3	4.4 ± 1.5
	1 位专家 vs. 1 位专家	no	8.8 ± 0.8	9.4 ± 0.9
		yes	7.9 ± 0.8	5.4 ± 1.2
			RGB	HSV
ca1	1 位专家 vs. 10 位用户	no	53.3 ± 2.6	*65.1 ± 2.9*
		yes	49.0 ± 2.6	64.1 ± 3.6
	1 位专家 vs. 1 位专家	no	76.7 ± 4.8	84.6 ± 4.9
		yes	68.3 ± 3.9	82.3 ± 4.1
ca2	1 位专家 vs. 10 位用户	no	39.1 ± 2.0	*42.6 ± 1.7*
		yes	35.4 ± 3.6	43.5 ± 3.6
	1 位专家 vs. 1 位专家	no	65.5 ± 5.8	74.2 ± 5.7
		yes	61.4 ± 4.9	70.4 ± 5.3

相比触觉数据，它们更容易受到环境噪声的影响；触觉数据同样可能包含特征性的低频分量。通过采用与基线相同的训练测试划分（1 名专家 + 9 用户对 1 名用户）及其交叉验证方法（无跨用户技巧），我们验证了关于波动来源的假设。这一更简单的分类问题取得了非常高的准确率（99.1% ± 0.9%），且与基线相比具有相近的用户差异。这仍然是跨用户训练-测试设置，而将我们的跨用户补偿应用于相同的五个信息源后，性能进一步提升（99.4% ± 0.8%）。总而言之，分类性能和鲁棒性能够受益于包含更多用户数据的更大数据库，尽管这会带来更高的计算成本。然而，这些结果也证实，我们的跨用户补偿策略对于仅由一名专家采集数据的数据库同样非常高效。

高效的跨用户补偿： 跨用户补偿消融研究（表 I）表明，移除我们的跨用户技巧会使所提出的完整分类器的均值准确率降低近 9%（降至 88.7%），同时使跨用户标准差增加 4.6%（升至 6.7%）。总体而言，核双样本检验受益于使用两个分布的全部统计矩进行比较。然而，对分布的均值进行对齐能够显著提升表面识别性能，并使其在不同用户之间更加稳健。我们认为，这一跨用户技巧之所以有效，是因为会话相关效应和用户相关效应集中在所提取分布的均值中，而高阶统计矩主要传达判别性表面信息。在针对各个信息源开展的消融研究（表 II）中，我们报告了在 1 对 10 设置下，将跨用户技巧与频谱采样结合时，所有振动均获得了更好的性能结果。跨用户技巧对时域振动采样的影响并不明显。相比之下，对摩擦力（fr）进行时域采样时，应用跨用户技巧可使分类性能提升 6%。

对于两幅图像，应用跨用户技巧并不会显著降低识别性能。这可能是因为用于采集这些图像的智能手机摄像系统会通过硬件和软件自动减弱某些与用户及会话相关的

图像效应。尽管如此，将跨用户技巧应用于两幅图像以及五个时间序列信息源时，我们仍获得了 97.1% ± 2.2% 的高分类准确率。虽然提出的补偿技巧主要用于跨用户分类，但在专家试次分类（一对一设置）中使用该技巧，会显著降低 ac（谱）、fr 和 ir 这三个信息源的识别率（表 II）。因此，在对这些分布均值进行对齐的过程中，有关表面身份的有用信息会丢失。相比之下，将跨用户技巧应用于 mi1 和 ac（时间）这两个信息源时，我们观察到专家试次分类性能有所提升。敲击声（mi1）是唯一一个专家试次分类识别率远低于 1 vs. 10 训练测试划分的来源。这表明专家的十次敲击试次差异很大，导致我们在一对一设置下的单源分类器难以分类。同时，正如图 6(c) 所示，这组多样化的敲击数据显著有助于其他信息源实现跨用户识别。

缺失的敲击加速度数据：我们无法将缺失的敲击加速度纳入分类器，因此无法与以往已发表方法的性能进行完美比较。敲击是一种探索性操作，能极大地帮助人识别表面 [50]。在双信息源分类中，结合上述多样化的专家试次，敲击声使另一信息源的识别准确率提高了 21% 至 50%（图 6(c)）。由于所有敲击声和加速度数据均来自相同的训练或测试交互，我们的分类器很可能会受益于测试阶段缺失的加速度信息源。预计会有这种提升，是因为额外的加速度数据代表了工具与表面交互的新瞬态信息，而且最大均值差异能够有效区分（新）数据差异。即使缺失的敲击加速度会降低多模态分类器的性能，其透明的架构也能轻松揭示这一趋势；与若干机器学习和深度学习模型不同，无需使用另一组信息源重新训练。

二元金属信号：我们对二元金属录音应用核双样本检验，以便为所有信息源采用统一而简单的算法架构。根据给定的金属检测读数，只考虑数据库中的两个子组之一（导电材料或非导电材料），有助于识别任务并降低算法的计算成本。此外，我们对该信息源的所有 MMD 得分进行平移，以避免在几何平均框架中出现 MMD $= 0$。将所有 MMD 得分平移为正值也很重要，因为在某些情况下，不应将所有决策权都交给单一信息源，例如安全关键型应用。

纹理特定见解：我们可以进一步考察最优分类设置（$S = 8$）的性能。对于四种较难的表面类别（网格、石材、光泽材料和橡胶），我们正确分类了全部 360 个测试样本；而对于木质材料，仅出现了（2/130）次类别内混淆（图 6(a) 和 (b)）。这类坚硬的表面纹理会放大接触诱发信号，因此能传递更丰富、更多鲜明的信息。在为数不多的严重误分类（30/1080）中，有十次表面交互被误判为地毯或泡沫材料。在大多数这些误分类中，正确的表面都位列排名靠前的候选项之中，这提示我们也可以考虑 $k > 1$ 最近邻。闪光纸是常见的假阳性候选项（出现 8 次）；这类纹理上突发的粘滑动作可能会产生异常且分布不规则的接触信号，从而使分类器混淆。

互补表面信息：成对信息源的分类性能（图 6(c)）展现了图像与拖动加速度信号结合时出色的分类准确率（87.9%）。直观而言，接触式与非接触式传感器可能会提供互补的表面信息。如果这两个信息源（ca1+ac）最终能在识别流程中长期发挥作用，那么便可以轻松调整总体决策的信息源加权。例如，通过公式 (11) 提高仅这两个信息源的权重（$w_{ca1} = w_{ac} = 3$），我们便能使 LMT108表面数据集的平均准确率达到 98.5% ± 1.2%。

多用户表面分类的难点：据我们所知，目前还没有针对整个 LMT108数据集或任何其他多用户表面数据集进行表面分类训练的深度学习算法。现有参考方法 [6], [9], [30] 均只考虑来自一个人的数据，因此无法与我们的多用户表面识别任务进行比较；我们认为后者与实际应用更为相关。在此背景下，我们通过消融研究展示了，只需简单地移动分布均值，就能有效补偿时间序列数据中的伪影，并提高仅有一个人采集的有限训练数据条件下的性能。

为了了解我们的分类器与最相近的深度学习参考方法相比表现如何，可以看到，我们的图像加触觉（ca1+ac）分类器在 LMT108数据集的一对一设置中取得了非常高的性能（97.1% ± 1.2%）；这一性能与一种深度学习分类器相当，后者使用图像与加速度输入，对由一位专家采集的 69 种表面的数据进行分类（属于 LMT69数据集）。请注意，根据相关消融研究的发现，我们并未采用跨用户补偿来获得这些结果。使用大量训练集（例如通过标准交叉验证）报告多模态表面识别性能，也有助于未来与深度学习网络进行比较。因此，我们报告了 LMT108数据集十对一设置下，仅使用标准图像时准确率为 88.9% ± 3.7%，使用图像加触觉（ca1+ac）输入时准确率为 96.9% ± 1.4%。在这一多用户设置中，我们仅对拖动加速度信号应用了跨用户技巧。总体而言，与深度学习分类器相比，我们的方法所需数据更少、可解释性更高，训练工作也简单得多。

核函数选择：平方指数核是一种常用选择，其超参数启发式方法已得到充分验证，而且无需进一步进行核函数调优，这一点很有优势。该核函数在 LMT108 表面图像和时间序列上表现良好。我们没有系统地研究其他核函数，但在研究初期尝试过几种其他选择。

例如，拉普拉斯核结合中位数启发式方法似乎能取得相近的识别性能。然而，其他一些核函数需要进行相当复杂的超参数优化，而我们希望避免这种情况。

VI. 结论

自动化解决方案在大数据计划中日益普及，在表面识别和物体识别等新兴领域尤为重要。机器学习和深度学习已广泛用于分类任务，但这些方法需要专家知识、启发式方法和／或特征工程。为避免这些流程，我们开发了一种自动分类框架，利用核两样本检验同时分析图像和时间序列。该方法易于实现，其原理是比较概率分布空间中的数据，并自动量化描述性差异。我们提出的算法架构清晰易懂、前后一致，为确定哪些传感器信息能够提升分类性能提供了一种简洁的方法。针对包含 108 个类别的表面识别任务，我们证明，该方法的准确率比采用专家设计特征的竞争性多模态基线分类器高 6%。在标准算法设置下，我们仅使用较少的传感器测量数据和训练数据，并进行少量数据处理，便实现了这一性能。我们还讨论了如何进一步提升多模态分类器的识别性能，实践者可根据具体应用场景采用并调整该方法（例如核函数、源权重）。利用更精确的传感系统采集新的表面交互数据，或增加测量数据（例如工具速度、工具角度），都可能为表面分类带来更多见解。在训练库中使用紧凑且具代表性的类内集合，有助于提高可扩展性。未来的工作还可以包括应用最大均值差异度量的改进版本，以提升检验功效并降低计算成本。此外，若能在信息源更少、训练数据稀疏的情况下实现相当的性能，将很有意义，也可借此与现有深度学习网络的需求进行对比；实现这一目标可能需要在训练阶段投入更多精力。我们还设想将此方法用于聚类等其他学习任务。

参考文献

[1] B. A. Richardson 和 K. J. Kuchenbecker，“学习预测触觉形容词的感知分布”，《神经机器人学前沿》，第 13 卷，第 116 页，2020年。[2] H. Liu、F. Wang、F. Sun 和 B. Fang，“利用弱配对跨模态学习检索表面材料”，*IEEE*《自动化科学与工程汇刊》，第 16 卷，第 2 期，第 781–791 页，2018年。[3] W. Zheng、H. Liu、B. Wang 和 F. Sun，“新物体的跨模态材料感知：一种深度对抗学习方法”，*IEEE*《自动化科学与工程汇刊》，第 17 卷，第 2 期，第 697–707 页，2019年。[4] H. Liu、D. Guo、X. Zhang、W. Zhu、B. Fang 和 F. Sun，“面向视障人士的图像到触觉跨模态感知”，*IEEE*《自动化科学与工程汇刊》，第 18 卷，第 2 期，第 521–529 页，2020年。[5] A. Burka、S. Hu、S. Helgeson、S. Krishnan、Y. Gao、L. A. Hendricks、T. Darrell 和 K. J. Kuchenbecker，“Proton：一种用于机器人学习表面属性的视觉-触觉数据采集系统”，载于*IEEE*智能系统多传感器融合与集成国际会议论文集（*MFI 2016*），2016年，第 58–65 页。

[6] M. Strese，《触觉材料采集、建模与显示》，博士论文，慕尼黑工业大学，2021年。[7] M. Strese、L. Brudermueller、J. Kirsch 和 E. Steinbach，《受人类探索程序启发的触觉材料分析与分类》，IEEE《触觉汇刊》，卷 13，期 2，页 404–424，2019年。[8] R. Geirhos、J.-H. Jacobsen、C. Michaelis、R. Zemel、W. Brendel、M. Bethge 和 F. A. Wichmann，《深度神经网络中的捷径学习》，《自然·机器智能》，卷 2，期 11，页 665–673，2020年。[9] H. Zheng、L. Fang、M. Ji、M. Strese、Y. ¨Ozer 和 E. Steinbach，《利用触觉和视觉信息进行表面材料分类的深度学习》，IEEE多媒体汇刊，卷 18，期 12，页 2407–2416，2016年。[10] A. Gretton、K. M. Borgwardt、M. J. Rasch、B. Sch¨olkopf 和 A. Smola，《核双样本检验》，《机器学习研究杂志》，卷 13，期 1，页 723–773，2012年。[11] F. Solowjow、D. Baumann、C. Fiedler、A. Jocham、T. Seel 和 S. Trimpe，《用于动力系统的核双样本检验》，arXiv预印本 arXiv:2004.11098，2020年。[12] M. Strese、Y. Boeck 和 E. Steinbach，《基于内容的表面材料检索》，载于 IEEE世界触觉大会（WHC）论文集，2017年，页 352–357。[13] M. Holliins、R. Faldowski、S. Rao 和 F. Young，《触觉表面纹理的感知维度：多维尺度分析》，《知觉与心理物理学》，卷 54，期 6，页 697–705，1993年。[14] W. M. B. Tiest 和 A. M. Kappers，《通过多维尺度分析及粗糙度和可压缩性的物理测量分析材料的触觉感知》，《心理学学报》，卷 121，期 1，页 1–20，2006年。[15] S. Okamoto、H. Nagano 和 Y. Yamada，《纹理触觉感知的心理物理维度》，IEEE《触觉汇刊》，卷 6，期 1，页 81–93，2012年。[16] J. A. Fishel、G. E. Loeb、B. Matulevich 和 R. Davoodi，《一种基于物体可能给人的触感来测量其触觉属性的方法及应用》，2016年1月28日，美国专利申请 14/796,647。[17] M. Strese、C. Schuwerk、A. Iepure 和 E. Steinbach，《基于多模态特征的表面材料分类》，IEEE《触觉汇刊》，卷 10，期 2，页 226–239，2016年。[18] Y. Shao、V. Hayward 和 Y. Visell，《全手触觉交互过程中的皮肤振动空间模式》，《美国国家科学院院刊》，卷 113，期 15，页 4188–4193，2016年。[19] Y. Liu、S. Lu 和 H. Culbertson，《通过音频－触觉跨模态一致性进行纹理分类》，载于 2022年 IEEE触觉研讨会，2022年，页 1–7。[20] A. Burka 和 K. J. Kuchenbecker，《处理触觉表面分类中的扫描时间参数》，载于 IEEE世界触觉大会（WHC）论文集，2017年，页 424–429。[21] J. Sinapov、V. Sukhoy、R. Sahai 和 A. Stoytchev，《人形机器人对表面的振动触觉识别与分类》，IEEE机器人学汇刊，卷 27，期 3，页 488–497，2011年。[22] N. Jamali 和 C. Sammut，《多数投票：利用表面纹理进行触觉传感材料分类》，IEEE机器人学汇刊，卷 27，期 3，页 508–521，2011年。[23] J. A. Fishel 和 G. E. Loeb，《用于纹理智能识别的贝叶斯探索》，《神经机器人学前沿》，卷 6，页 4，2012年。[24] V. Chu、I. McMahon、L. Riano、C. G. McDonald、Q. He、J. M. Perez-Tejada、M. Arrigo、T. Darrell 和 K. J. Kuchenbecker，《通过物理交互学习触觉形容词的机器人学习》，《机器人与自主系统》，卷 63，页 279–292，2015年。[25] M. Kaboli 和 G. Cheng，《利用人工机器人皮肤，根据纹理属性区分物体的鲁棒触觉描述符》，IEEE机器人学汇刊，卷 34，期 4，页 985–1003，2018年。[26] R. Bello-Cerezo、F. Bianconi、F. Di Maria、P. Napoletano 和 F. Smeraldi，《理想与现实条件下用于彩色纹理分类的手工设计图像描述符与现成CNN特征的比较评估》，《应用科学》，卷 9，期 4，页 738，2019年。[27] A. Krizhevsky、I. Sutskever 和 G. E. Hinton，《利用深度卷积神经网络进行ImageNet分类》，《神经信息处理系统进展》，卷 25，页 1097–1105，2012年。[28] L. Liu、J. Chen、P. Fieguth、G. Zhao、R. Chellappa 和 M. Pietik¨ainen，《从词袋模型到卷积神经网络：纹理分类中纹理表示的二十年发展》，《国际计算机视觉杂志》，卷 127，期 1，页 74–109，2019年。[29] M. Cimpoi、S. Maji 和 A. Vedaldi，《用于纹理识别与分割的深度滤波器组》，载于 IEEE计算机视觉与模式识别会议论文集，2015年，页 3828–3836。

[30] J. B. Joolee、M. A. Uddin 和 S. Jeon，《基于深度多模型融合网络的触觉数据真实物体触觉理解》，《应用智能》，页 1–16，2022年。[31] Y. Gao、L. A. Hendricks、K. J. Kuchenbecker 和 T. Darrell，《利用视觉与触觉数据进行触觉理解的深度学习》，载于 IEEE国际机器人与自动化会议（ICRA）论文集，2016年，页 536–543。[32] J. Wei、C. Shaowei、P. Hao、J. Hu、S. Wang 和 Z. Lou，《多模态未知表面材料分类及其在物理推理中的应用》，IEEE《工业信息学汇刊》，2021年。[33] B. Sch¨olkopf 和 A. J. Smola，《核方法学习：支持向量机、正则化、优化及其他》，麻省理工学院出版社，2018年。[34] K. Muandet、K. Fukumizu、B. Sriperumbudur 和 B. Sch¨olkopf，《分布的核均值嵌入：综述及进展》，《机器学习基础与趋势》，卷 10，期 1-2，页 1–141，2017年。[35] M. Long、Y. Cao、J. Wang 和 M. Jordan，《利用深度适应网络学习可迁移特征》，载于国际机器学习会议论文集，2015年，页 97–105。[36] I. Tolstikhin、O. Bousquet、S. Gelly 和 B. Sch¨olkopf，《Wasserstein自编码器》，载于国际学习表征会议（ICLR 2018）论文集，2018年5月。[37] A. Rozantsev、M. Salzmann 和 P. Fua，《深度域适应：超越共享权重》，IEEE《模式分析与机器智能汇刊》，卷 41，期 4，页 801–814，2018年。[38] L. C. Vankadara、S. Bordt、U. Luxburg 和 D. Ghoshdastidar，《非参数混合模型下基于核的聚类的恢复保证》，载于国际人工智能与统计会议论文集。PMLR，2021年，页 3817–3825。[39] G. Doran、K. Muandet、K. Zhang 和 B. Sch¨olkopf，《一种基于置换的核条件独立性检验》，载于《人工智能中的不确定性》，2014年，页 132–141。[40] K. Chwialkowski 和 A. Gretton，《随机过程的核独立性检验》，载于国际机器学习会议论文集，2014年，页 1422–1430。[41] J. R. Lloyd 和 Z. Ghahramani，《使用核双样本检验进行统计模型批评》，《神经信息处理系统进展》，卷 28，页 829–837，2015年。[42] S. Lu、M. Zheng、M. C. Fontaine、S. Nikolaidis 和 H. Culbertson，《通过交互式生成与搜索进行偏好驱动的纹理建模》，IEEE《触觉汇刊》，卷 15，期 3，页 508–520，2022年。[43] B. Khojasteh、M. Janko 和 Y. Visell，《手指与纹理表面之间滑动摩擦动力学的复杂性、速率与尺度》，《科学报告》，卷 8，期 1，页 1–10，2018 年。[44] D. Sejdinovic 、 B. Sriperumbudur 、 A. Gretton 和 K. Fukumizu，《假设检验中基于距离的统计量与基于 RKHS的统计量的等价性》，《统计年刊》，页 2263–2291，2013年。[45] A. Gretton、K. Fukumizu、C. H. Teo、L. Song、B. Sch¨olkopf 和 A. J. Smola，《一种核独立性统计检验》，载于《神经信息处理系统进展》，卷 20，2007年，页 585–592。[46] H. Culbertson、J. Unwin、B. E. Goodman 和 K. J. Kuchenbecker，《从不受约束的工具—表面交互中生成触觉纹理模型》，载于 IEEE世界触觉大会（WHC）论文集，2013年，页 295–300。[47] M. Janko、R. Primerano 和 Y. Visell，《主动触摸过程中手指与纹理表面之间的摩擦力》，IEEE《触觉汇刊》，卷 9，期 2，页 221–232，2015年。[48] P. Flach，《机器学习：让数据变得有意义的算法之艺术与科学》，剑桥大学出版社，2012年。[49] M. Mohri、A. Rostamizadeh 和 A. Talwalkar，《机器学习基础》，麻省理工学院出版社，2018年。[50] R. H. LaMotte，《借助工具进行软硬辨别》，《神经生理学杂志》，卷 83，期 4，页 1777–1786，2000年。

附录 A 软件实现

我们使用开源分类处理流程**(4)**测试 LMT108 表面数据集**(1)**，采用未经修改的标准核双样本检验**(2)** 和希尔伯特-施密特独立性准则**(3)**实现。为确保结果可复现，我们的开源代码包含随机数生成器的默认设置。原始来源链接如下：**(1)**：https://zeus.lmt.ei.tum.de/downloads/texture**(2)**：http://www.gatsby.ucl.ac.uk/~gretton/mmd/mmd.htm

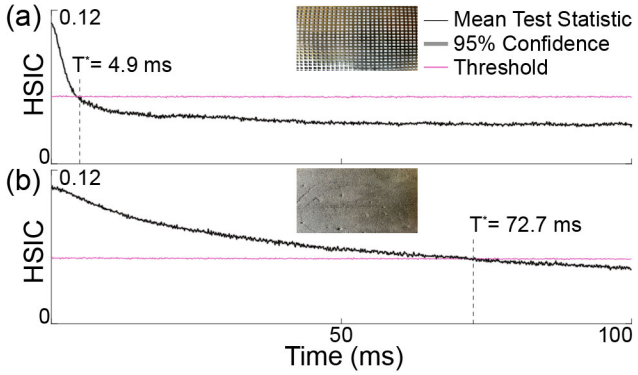


图 7. 对于测得的红外光反射率，HSIC（$R = 500$, $\alpha = 0.05$, $t_{1,\max} = 1$ 秒）随时间的变化，测量对象为 (a) 铝网和 (b) EPDM 泡沫（数据来自 LMT108 训练数据 [12]）。与平坦的吸收性泡沫表面相比，金属网的红外传感器样本混合速度约快 15 倍。

(3)：http://people.kyb.tuebingen.mpg.de/arthur/indep.htm**(4)**：论文发表后公开代码

附录 B 表面信号的混合特性

消失自相关时间 T^* 确保样本近乎相互独立，从而适用于核双样本检验。此外，它还具有物理学上的解释，并能帮助洞察某些系统属性。

对于表面信号，这些混合特性体现了针对给定信息源 s，与特定表面 c 交互所产生的不同试次具有多高的丰富度。例如，红外反射传感器的时间序列测量数据在两种不同表面上呈现出有趣的希尔伯特-施密特独立性准则结果（图 7）。铝网表面只需约 5 毫秒，样本之间便能获得新的不相关信息；而泡沫表面的混合速度则慢约 15 倍。三维多孔网状表面能产生丰富的信号，而平坦黑色泡沫表面会大量吸收红外信号，因而信号丰富度较低。

从更广泛的角度来看，这些表面特定混合特性为获取给定信息源的新测量数据提供了节省时间的指导。可以针对每个表面单独调整测量时间，以避免自相关并获得丰富的独立数据。

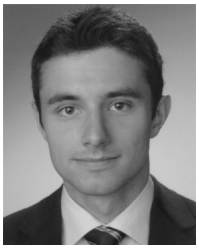
致谢

作者感谢 Strese 博士和 Steinbach 教授对 LMT 数据库所作的有益说明，感谢 Ben Richardson 对手稿提出的宝贵反馈，并感谢国际马克斯·普朗克智能系统研究生院（IMPRS-IS）对 Behnam Khojasteh 和 Friedrich Solowjow 的支持。



Behnam Khojasteh 分别于2014年和2017年在德国汉堡的汉堡工业大学获得机械工程和机电一体化专业的理学学士学位和理学硕士学位。在硕士学习期间，他曾在美国加利福尼亚大学圣塔芭芭拉分校电气与计算机工程系担任访问研究员。Khojasteh 目前是德国斯图加特马克斯·普朗克智能系统研究所触觉智能系的博士生，同时也是国际马克斯·普朗克

国际马克斯·普朗克智能系统研究生院。他在汽车和航空航天领域拥有两年多的行业经验。



Friedrich Solowjow 分别于2014年和2015年获得波恩大学数学和经济学学士学位，并于2017年获得该校数学硕士学位。2022年，他通过国际马克斯·普朗克智能系统研究生院（国际马克斯·普朗克智能系统研究生院）获得图宾根大学自然科学博士学位（Dr. rer. nat.）。目前，他是德国亚琛工业大学机械工程数据科学研究所的高级研究员（学术顾问）。

他的主要研究兴趣在于系统与控制理论和机器学习的交叉领域。



Sebastian Trimpe 于2005年和2007年分别获得汉堡工业大学（德国汉堡）通用工程科学学士学位和电气工程硕士学位（Dipl.-Ing.），并于2013年获得苏黎世联邦理工学院（瑞士苏黎世）机械工程博士学位（Dr.sc.）。自2020年起，他担任德国亚琛工业大学正教授，并领导机械工程数据科学研究所。此前，他曾在马克斯

普朗克智能系统研究所（德国斯图加特和图宾根）担任独立研究组负责人。他的主要研究兴趣包括系统与控制理论、机器学习、网络化系统和机器人学。Trimpe博士曾获多项荣誉，包括每三年颁发一次的IFAC世界大会互动论文奖（2011年）、克劳斯·奇拉公众科学理解成就奖（2014年）、国际信息物理系统会议最佳论文奖（2019年），以及埃瓦尔德·马夸特基金会颁发的控制工程创新未来奖（2020年）。



Katherine J. Kuchenbecker 于2000年、2002年和2006年分别获得美国斯坦福大学机械工程学士、硕士和博士学位。她曾在约翰斯·霍普金斯大学从事博士后研究，并于2007年至2016年在宾夕法尼亚大学GRASP实验室担任教职。自2017年起，她一直担任德国斯图加特马克斯·普朗克智能系统研究所的主任。她的研究融合了机器人学与人机交互，重点关注触觉技术、远程操作、人机物理交互、

动作、触觉传感和医疗应用。她于2012年发表了一场关于触觉技术的TEDYouth演讲，并荣获2009年NSF职业奖、2012年IEEE RAS学术早期职业奖、2014年宾夕法尼亚大学林德巴克杰出教学奖、2022年当选IEEE会士，以及多个最佳论文、海报、演示和审稿人奖。她于2015年至2017年共同主持IEEE RAS触觉技术委员会，并于2016年和2018年共同主持IEEE触觉研讨会。