

指导教师：         杨涛        

提交时间：         2016/3/17        

# CVPR2015 Paper

## Translation

No:         1        

姓名：         郑在晖        

学号：         2013302611        

班号：         10011306        



## 一种稳定的多核拓扑机器学习

### 摘要

拓扑数据分析提供了大量有价值的信息来研究视觉问题。然而，到目前为止，我们缺少与大众普遍接受的基于核学习技术相关联的一种合理的理论解释，例如核支持向量机或者核 PCA。在这项工作中，我们建立一种连接，这种连接是通过设计一个多核持久性的图表，一种稳定的数据拓扑特征的概要表示。我们发现，这种内核是正定而且能够证明其稳定性相对 1-瓦瑟斯坦距离。实验对三维形状分类/检索的 2 个基准数据集和纹理识别表现出的性能比基于最近介绍持久性风景的另一种方法更好。

### 引言

在许多计算机视觉问题，数据（例如，图像，网格，点云等）为了从链中提取有用的信息，用于解决高层推理任务，如识别，检测或分割而通过复杂的管道处理。所提取的信息，可能会以低级别外观描述的形式，例如，SIFT[20]，或更高级别的性质，例如，活动在深度卷积网络的特殊层[18]。在识别问题，例如，固定输入数据给判别分类器，例如目前广受欢迎的支持向量机（SVM），基于内核学习技术。

虽然已经在提取和编码区别信息上取得实质性进展，直到最近才有人开始调查拓扑结构的数据作为信

息的额外来源，随着拓扑数据分析（TDA）[4]，计算机工具能够高效的识别拓扑结构已成为现实。从那时以来，一些作者已经证明 TDA 的方法可以捕获其他方法往往不能揭露出的特性数据 CF [26, 19]。

沿着这些思路，用 TDA 来研究持续的同源性[11]是一种特别受欢迎的方法，因为它在多尺度法捕捉了拓扑特征开始和消亡的时间，例如，连接部件，孔等。这些信息以摘要方式表示在连续图，即平面上的一个多重点集。持久性的同源性的主要特征是它的稳定性：输入数据细微的变化将会引起相关联连续图[10]的瓦瑟斯坦距离的微小变化。考虑到拓扑信息的离散型，存在这样良性的结论或许不足为奇的。

注意到连续图和瓦瑟斯坦距离只形成一个度量空间。因而这是不可能在大类机器学习技术直接采用在持久性的同源性中，因为这要求希尔伯特空间结构，支持向量机(SVM)或 PCA。这个障碍通常是在包含数据领域定义一个内核函数，进而暗中限定了希尔伯特空间结构。虽然瓦瑟斯坦距离本身不会自然引出一个有效的内核（见补充材料详情），我们发现，这可能规定一种核持久性图表，它是 wrt 稳定，即 1-瓦瑟斯坦距离。这是这篇论文主要的贡献。

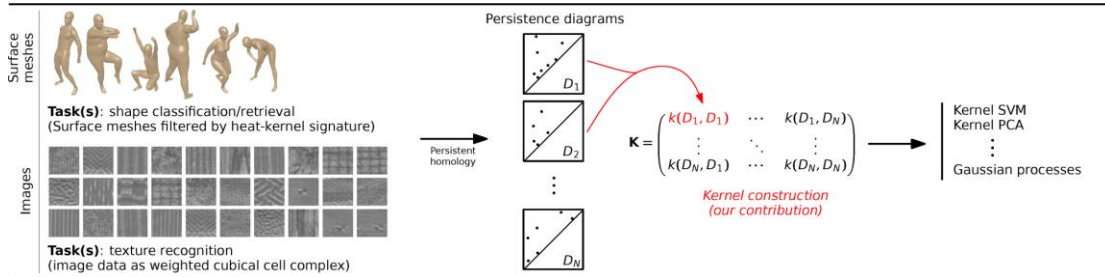


图 1：视觉数据（例如，在表面网格，纹理，等等功能）被使用持久的同源性[11]进行分析。

大致说来，持续的同源性在持续性的形式捕获的拓扑特征（例如，连接组件或空穴）出生/死亡时间图。我们的贡献是定义一个内核，持久性的图表，使理论上在计算机视觉界流行的框架内合理使用这些总结来表示基于内核的学习技术。

贡献：我们提出了一个（正定）多尺度内核持久图（见图 1）。这个内核通过 L2 值的特征图谱定义，基于尺度空间理论[16]。我们发现我们的特征图谱的利普希茨连续条件和 1 - 瓦瑟斯坦距离有关，从而能够保持具有稳定性能的持久性的同源性。我们的内核的尺度参数让其自身对噪音具有鲁棒性，并且可以被动的调准数据。我们详细的调查分析内核的理论性能并且能够证明可以适用于形状分类/检索和纹理识别基准。

## 2. 相关工作

利用拓扑信息的方法对计算机视觉或医学图像分析大致可以分为两类。在第一类中，我们确定以前的工作是直接利用拓扑信息处理一些特殊问题，如拓扑分割。在第二类中，我们确定处理间接使用拓扑信息的方法。换句话说，关于拓扑特征的信息被用作一些机器学习算法的输入。

作为第一类的代表，Skraba 等

[26]，适应于基于持久性聚类[6]的方法，在三维形状的曲面网格分割上，通过由持久性图中的拓扑信息驱动。Gao 等，[12]使用持久性的信息来恢复把柄，即，拓扑周期，在已经分割的左心室，从中提取计算机断层扫描图像。在不同的细胞分裂组织，Chen 等。[7]打算直接合并拓扑约束作为随意基本分割模式。

在接近第二类，Chung 等 [8]和 Pachauri 等人。[22]调查分析皮质厚度测量人类的大脑皮层的 3 维曲面网格的问题，为了研究发育和神经障碍。相反，[26]，持久性信息没有直接使用，而是作为一个描述符供应给判别分类器。为了区分正常患者和阿尔茨海默氏病/自闭症。然而，训练分类器的步骤中拓扑信息以一种特别的方式输入。在[22]例如，在一个规则网格，持久性图首先被栅格化，然后计算内核密度估计，最后向量化离散概率

密度函数被用作特征向量来训练采用标准内核  $R$  的 SVM。然而，目前还是不清楚内核对现有指标在感应距离，（例如，瓶颈(bottleneck)或瓦瑟斯坦距离），以及稳定性受到影响的性能如何。由 Li [19] 等人最近提出的一种方法是直接使用固定距离的持久性图来识别。除了瓶颈(bottleneck)和瓦瑟斯坦距离，作者在实验中采用的持久性景观 [3]，和相应距离。实验结果表明当加上传统的 bag-off-eature 方法处理时会显示持久信息的互补性。虽然我们以秒为单位研究。5. 2 [19] 的启发，我们首先注意到内核的发展；明确结合其他方法。

为了能够在机器学习体系中使用持久性信息，阿德科克等。 [1] 提出了用使由代数几何和不变的理论获得的特征向量比较持久性图。特点是使用代数函数定义持久图生死价值。

从概念的观点，Bubenik 的概念（持久性景观 [3]）可能是最接近我们的，作为另一种持久图的特征图谱。虽然持久性景观没有明确的使用在机器学习算法，我们将这种概念结

合到我们的工作部分。5. 1，表明他们事实上承认有效正定核的定义。此外，无论是持久性的景观，还是我们的方法都计算的表示替代瓶颈(bottleneck)或瓦瑟斯坦的距离，这两者都需要匹配问题的解决方案。

### 3. 背景

首先，我们从持续的同源性回顾一些基本概念和结果，这将会对我们的工作有意义。

持久性图表（Persistence diagrams）：持久性图是一个简洁描述不断变化序列形状的拓扑变化，被称为滤波。特别是，随着形状的变化，不同尺寸的孔中（即，部件，隧道，空隙等之间的间隙）可能会出现和消失。直观地，一个  $k$  尺度孔，产生在  $a$  时间，填充外  $b$  时间，产生了一个点  $(A, B)$  在第  $k$  个持久图。因此，一个持久化图在  $R^2$  是多重点集。从形式上看，持久性图是利用代数拓扑，即所谓的同源性，的一个标准概念定义；见 [11] 的细节。

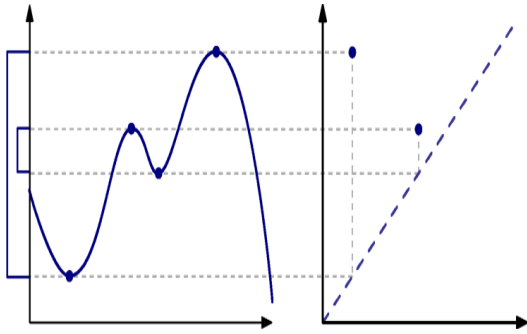


图 2：一个函数  $R \rightarrow R$ (左)和它的  $0^{\text{th}}$ 持久性图(右)，局部极小中在相应的水平子集产生一个连通分量。而局部最大值合并连接组成部分。出生和死亡的配对在持久图上所示。

注意，不是每一个孔都会在过滤中消失。这样的孔的基本特征发生改变和表现在图中的  $(a, \infty)$  点集中。因此，在过滤后捕获最后的拓扑形状。在目前的工作，我们不考虑这些特征作为持久性图的一部分。此外，所有的持久性图将被假定为有限的，常是如此来从数据计算持久性图。

**函数过滤**：一个得到过滤标准的方法是考虑水平子集  $f^{-1}(-\infty, t]$  在函数  $f: \Omega \rightarrow R$  定义一些域  $\Omega, t \in R$ . 为了易于明显的看出水平子集来自过滤，用  $t$  参数表示。用  $Df$  表示持久性图结果；看参照图 2 的图示。

作为一个例子，考虑一个灰度图像，其中  $Q$  是图像中的矩阵域和  $f$  是在域中的灰度值（即，在特定的像素）。因此，一个水平子集包含  $\Omega$  中的所有像素点，并且值趋向于某一临界值  $t$ 。

另一个例子是在三角网格  $Q$  的一个分段线性函数，例如广受欢迎热核信号 [27]。然而，另一种常用的用法是由一个点云  $p$  嵌入在  $R^n$ , 通过考虑距离函数  $d_P(x) = \min_{y \in P} \|x - y\|$  在  $\Omega = R^n$ . 该函数的水平子集是围绕  $P$  上每个点的集合，就计算而言，它们通常由被称为  $\alpha$  形状的等价物替代。

**稳定性**：在持久图  $Df$  的函数  $f$  的一个重要方面是相对扰动  $f$  的稳定性。事实上，只有稳定保证才可以推断出函数  $f$  在噪音的干扰下在持久性图中的信息。

形式上，我们认为  $f \rightarrow Df$  作为地图度量空间并定义为稳定性为李普希茨连续性。这个要求选择的度量既要是函数的集合同时也是持久性图的集合。对于函数，通常采用  $L^\infty$  度量尺度。

对于持久性的图采用自然度量，称为瓶颈(bottleneck)距离。宽松的讲，两个图表的距离是以通过任何两个相应的点的最大距离的最小参数表示，在两图之间的所有双射上。形式上，设  $F$  和  $G$  代表两种持久性图，每个图通过对角线加入可数无穷重根的点  $(t, t)$  扩张。这个瓶颈(bottleneck)距离为

$$d_B(F, G) = \inf_{\mu} \sup_{x \in F} \|x - \mu(x)\|_{\infty} \quad (3)$$

(1)

其中， $\mu$  范围包括所有双射，即从  $F$  中的一小部分到  $G$  中的单个元素。如由科恩斯坦纳等。 [9]，持久性的图表对于瓶颈 (bottleneck) 距离相对于稳定。

瓶颈 (bottleneck) 距离嵌入到更一般级别的距离，称为瓦瑟斯坦 (Wasserstein) 距离。对于任何正实数  $p$ ，在  $p$ -瓦瑟斯坦距离是

$$d_{w,p}(F, G) = \left( \inf_{\mu} \sum_{x \in F} \|x - \mu(x)\|^p \right)^{1/p}$$

其中， $\mu$  范围包括所有双射，即从  $F$  中的一小部分到  $G$  中的单个元素。注意取极限  $p \rightarrow \infty$  域在瓶颈 (bottleneck) 距离，并且我们定义  $d_{w, \infty} = d_B$ 。我们有以下的结果边界  $p$ -Wasserstein 距离在  $L^\infty$  距离：

**定理1** (科恩-Steiner等[10])。假定  $X$  为压缩 triangulable 度量空间，使得对于  $L$ -Lipschitz 函数  $f$  在  $X$  上，并且  $k \geq 1$ ， $k$  度达到定值  $\Sigma(b, d) \in D_f(d-b)k$  有恒定上界  $C$ 。设  $f, g$  为在  $X$  中的两个  $L$ -Lipschitz 分段线性函数，并且对于所有的  $p \geq k$ 。

$$d_{w,p}(D_f, D_g) \leq (LC)^{1/p} \|f-g\|$$

我们注意到，严格来说，从某种意义上说，对于利普希茨连续 (Lipschitz continuity) 这不是一个稳定性结果，因为它只能创建 Holder 连续性。此外，当  $p=1$  时对于  $p$ -Wasserstein 距离它只给出了一个恒定上限。

**内核：**给定一组  $X$ ，一个函数  $K: X * X \rightarrow R$  为一个内核，如果存在一个 Hilbert 空间  $H$ ，称为特征空间，和一个地图  $\Phi: X \rightarrow H$ ，称为特征图，使得  $K(x, y) = \langle \Phi(x), \Phi(y) \rangle_H$ ，对于所有的  $x, y \in X$ 。相等于， $k$  为一个内核，如果是对称正定 [24]。核心程序允许在一个 Hilbert 空间去应用机器学习算法用来表示一般的设置，例如，字符串，图表，或者是我们的实例，持久性的图表。

内核在  $x$  上引起一个变度量  $dk(x, y) = (k(x, x) + k(y, y) - 2k(x, y))^{1/2}$ ，也就是在特征空间距离为  $\|\Phi(x) - \Phi(y)\|_H$ 。我们称核  $k$  稳定 w.r.t.  $X$  上的一个度量  $d$ ，如果有一个常数  $C > 0$ ，使得  $dk(x, y) \leq Cd(x, y)$ ，对所有的  $x, y \in X$ 。注意，这等同于 Lipschitz 连续性。

内核的稳定性对于分类问题是特别有用的：假定存在一个分离超平面  $H$  对于边缘  $m$  的两类数据点。如果数据点波动在  $\epsilon < m/2$ ，然后  $H$  将会扰动这两类在边缘  $m - 2\epsilon$ 。

#### 4. 持久性尺度空间的内核

我们提出了一个稳定的多尺度内核  $k_\sigma$  从持久性的图  $D$  的自身中

1)。这个内核将通过定义特征图  $\Phi_\sigma : D \rightarrow L^2(\Omega)$ ，并且  $\Omega \subset \mathbb{R}^2$  表示靠近半平面上面的对角线。

为了激励定义的  $\Phi_\sigma$ ，我们指出持久性的图的集合，即在  $\mathbb{R}^2$  的多重点集，使不具有 Hilbert 空间结构本身。然而，持久性图  $D$  可被唯一表示为一个狄拉克分布'的总和，对于每一个在  $D$  上的点。狄拉克分布是希尔伯特空间  $H^2(\mathbb{R}^2)$  的功能[17, 第 7 章]，我们通过采用这种观点获得一个持久性图的规范希尔伯特空间结构。

不幸的是，在  $D$  上的感应度量没有考虑点的在图中的距离或者对角线的距离，因此持久性图对于扰动不具有鲁棒性。受尺度空间理论的启发[16]，我们解决这个问题通过使用狄拉克的总和增量作为用于热扩散问题的初始条件与上对角线狄里克雷边界条件。这个偏微分方程的解决方案是一个  $L^2(\Omega)$  函数对于任何选定的尺度参数  $\sigma > 0$ 。在以下段落中，我们将

- 1) 定义持久性尺度空间的内核包  $k_\sigma$ ，
- 2) 推导出一个简单的公式来评估  $k_\sigma$ ，以及
- 3) 证明  $k_\sigma$  w.r.t 的稳定性, 1-瓦瑟斯坦距离。

需要注意的是  $\Phi_\sigma(D)=0$  对于一些

**定义 1** 设  $\Omega = \{X = (X_1, X_2) \in \mathbb{R}^2: X_2 \geq X_1\}$  表示上面的对角线空间中，并

令  $\delta_X$  表示狄里克雷条件中心在点  $x$ 。

对于给定的持续性图  $D$ ，我们现在考虑解  $u: \Omega \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}, (x, t) \rightarrow u(x, t)$  偏微分方程 2

$$\Delta_x u = a_x u \quad \text{in } \Omega \times \mathbb{R}_{>0} \quad (4)$$

$$u = 0 \quad \text{on } a\Omega \times \mathbb{R}_{\geq 0} \quad (5)$$

$$u = \sum_{y \in D} \delta_y \quad \text{on } \Omega \times \{0\} \quad (6)$$

<sup>1</sup> 一个狄拉克分布的一个功能是计算在一个点的平稳。

<sup>2</sup> 由于初始条件 (6) 不是  $L^2(\Omega)$  函数，该方程

是在分布上有意义。为了严格处理存在量化和独特的解决方案，见[17, 第 7 章]。

特征图  $\Phi_\sigma : D \rightarrow L^2(\Omega)$  在刻度  $\sigma > 0$ ，在一个持久性图  $D$  上定义  $\Phi_\sigma(D) = u|_{t=\sigma}$ ，这图域的持续性的尺度空间核  $k_\sigma$  在  $D$  上为

$$K_\sigma(F, G) = \langle \Phi_\sigma(F), \Phi_\sigma(G) \rangle_{L^2(\Omega)} \quad (7)$$

$\sigma > 0$  意味着  $u=0$  在  $\Omega \times \{0\}$ ，这意味着

D 具有必须是空图。从解决方案运算子的线性度，现在得出  $\Phi \sigma$  是单射映射。

偏微分方程的解决方案可以由得  $\Omega$  扩展到  $R^2$  和更换

$$u = \sum_{y \in D} \delta_y \cdot \overline{\delta_y} \quad \text{在 } R^2 \times \{0\} \quad (8)$$

所得，其中  $\overline{y} = (b, a)$  为  $y = (a, b)$  的镜像在反映。它可以表明出限制扩展到  $\Omega$  域产生的原始方程的解决方案。它通过用高斯核心卷积初始条件 (8) 给出：

$$u(x, t) = \frac{1}{4\pi t} \sum_{y \in D} \exp\left(-\frac{\|x-y\|^2}{4t}\right) - \exp\left(-\frac{\|x-\overline{y}\|^2}{4t}\right). \quad (9)$$

对 U 使用这种封闭形式的解决方案，我们可以得出一个简单的表达式来明确评估内核：

$$k_\sigma(F, G) = \frac{1}{8\pi\sigma} \sum_{\substack{y \in F \\ z \in G}} \exp\left(-\frac{\|y-z\|^2}{8\sigma}\right) - \exp\left(-\frac{\|y-\overline{z}\|^2}{8\sigma}\right). \quad (10)$$

我们参考补充材料作为初步推导 (10) 和一个可视化的解决方案 (9)。需要注意的是，内核可以在  $0 < (|F| \cdot |G|)$  计算当  $|F|$  和  $|G|$  表示多重集合 F 和 G 的基数时候，各自 (或接近有界误差在  $0 < (|F| + |G|)$  使用快速高斯变换 [14])。

**定理 2** 内核  $k_\sigma$  是 1 - 瓦瑟斯坦稳定。

证明：证明 1 Wasserstein 的稳定性，我们表明特征图  $\Phi \sigma$  Lipschitz 连续性如下：

$$\|\Phi_\sigma(F) - \Phi_\sigma(G)\|_{L_2(\Omega)} \leq \frac{1}{2\sqrt{\pi\sigma}} d_{W,1}(F, G), \quad (11)$$

其中，F 和 G 表示增强与对角线上点后的持久图。需要注意的是扩张图表

在对角线上的点不会改变  $\Phi \sigma$  的值，

如从 (9) 中可以看出。通过定义 Wasserstein 的距离，对于任何  $\delta > d_{W,1}(F, G)$  存在一个在 F 和 G 之间的匹配  $\mu$  与  $\sum_{y \in F} \|y - \mu(y)\|_\infty \leq \delta$ 。写作  $N_y(X) = 1/(4\pi\sigma) \exp(-\|x-y\|^2/4\sigma)$ ，我们有

$$\|N_y - N_z\|_{L_2(\mathbb{R}^2)} = \frac{1}{\sqrt{4\pi\sigma}} \cdot \sqrt{1 - \exp\left(-\frac{\|y-z\|^2}{8\sigma}\right)}.$$

$\Phi \sigma$  的定义从 (9) 和  $N_y$ ，闵可夫斯基 (Minkowski) 的定义不等式，不等式  $1 - \exp(-\xi) \leq \xi$ ，和不等式

$$\|x\|_2 \leq (\sqrt{2}) \|x\|_\infty \quad \text{最后域为}$$



$$\begin{aligned}
 & \|\Phi_\sigma(F) - \Phi_\sigma(G)\|_{L_2(\Omega)} \\
 &= \left\| \sum_{y \in F} (N_y - \bar{N}_y) - (N_{\mu(y)} - \bar{N}_{\mu(y)}) \right\|_{L_2(\mathbb{R}^2)} \\
 &\leq \left\| \sum_{y \in F} (N_y - N_{\mu(y)}) \right\|_{L_2(\mathbb{R}^2)} + \left\| \sum_{y \in F} (\bar{N}_y - \bar{N}_{\mu(y)}) \right\|_{L_2(\mathbb{R}^2)} \\
 &= 2 \sum_{y \in F} \|N_y - N_{\mu(y)}\|_{L_2(\mathbb{R}^2)} \\
 &= \frac{1}{\sqrt{\pi}\sigma} \sum_{y \in F} \sqrt{1 - \exp\left(-\frac{\|y - \mu(y)\|^2}{8\sigma^2}\right)} \\
 &\leq \frac{1}{\sqrt{8\pi}\sigma} \sum_{y \in F} \|y - \mu(y)\| \leq \frac{1}{2\sqrt{\pi}\sigma} \delta
 \end{aligned}$$

任何  $\delta > d_{w,1}(F, G)$ ，并要求如下。

我们指的是左边 (11) 作为持久尺度空间距离  $d_k$ 。注意的右边 (11) 减少当  $\sigma$  增大，调整  $\sigma$  以抵消在输入数据中的噪声的影响，这会导致在  $d_{w,1}(F, G)$  增加。我们将在秒单位看到在 5.3 调  $\sigma$  为数据对在机器学习方法的整体性能上有利。

从定理 2 所产生的一个自然的问题是，是否我们的稳定性结果延伸到  $p$ -Wasserstein 的距离要  $p > 1$ ，为了回答这个问题，我们首先注意到，我们的内核是附加的：我们所说的核  $k$  在持久性的图表附加如果  $k(E \cup F, G) = K(E, G) + K(F, G)$  对于所有  $E, F, G \in D$ 。通过选择  $F = \emptyset$ ，我们可以看到，如果  $k$  是附加的则  $K(\emptyset, G) = 0$  对全部  $F, G \in D$ 。接下来的定理建立定理 2 就是有意义的添加核可以稳定 w. r. t.  $d_{w,p}$ 。当  $p > 1$ 。

定理 3 一个有意义的附加核  $k$  在持久性图中存在一个  $F \in D$  例如  $k(F,$

$F) > 0$ 。我们证明通过比较  $d_k(\cup_{i=1}^n F, \emptyset)$  的增长率和  $d_{w,p}(\cup_{i=1}^n F, \emptyset)$  w. r. t.  $n$ 。我们有

$$d_k\left(\bigcup_{i=1}^n F, \emptyset\right) = n \sqrt{k(F, F)}.$$

另一方面，

$$d_{w,p}\left(\bigcup_{i=1}^n F, \emptyset\right) = d_{w,p}(F, \emptyset) \cdot \begin{cases} \sqrt[n]{n} & \text{if } p < \infty, \\ 1 & \text{if } p = \infty. \end{cases}$$

因此， $d_k$  不能被  $C \cdot d_{w,p}$  约束，对以一个常量  $C > 0$ ，如果  $p > 1$ 。

## 5. 评估

为了在秒单位级评估提出的内核。4，我们调查与秒为单位持久性景观方法在概念上的差异。5.1，再考虑以秒为单位的形状分类/检索和纹理识别背景下的性能。5.2。

### 5.1 比较持久性景观方法 (persistence landscapes)

在 [3], Bubenik 推出持续性景观方法，是将持久性图表示在 Banach 空间  $L_p(\mathbb{R}^2)$  中。这种结构的主要目的是用于统计计算，使用矢量空间结构  $L_p$ 。对于  $p=2$ ，我们可以使用希尔伯特  $L_2(\mathbb{R}^2)$  的空间结构，构建一个类似的内核 (7)。对于这项工作的目的，我们把这个内核作为持久性景观内核  $k^L$ ，并定义

$\Phi^L: D \rightarrow L_2(\mathbb{R}^2)$  相应特征图。该内核感应距离由  $d_{kl}$  表示。Bubenik 表示稳定 w.r.t. 在 Wasserstein 加权模式下的距离，这对于  $p=2$  可以被概括为：

定理 4 (Bubenik[3])。对于任何两个持续性的图表  $F$  和  $G$ ，我们有

$$\|\Phi^L(F) - \Phi^L(G)\|_{L_2(\mathbb{R}^2)} \leq \inf_{\mu} \left( \sum_{y \in F} \text{pers}(y) \|y - \mu(y)\|_{\infty}^2 + \frac{2}{3} \|y - \mu(y)\|_{\infty}^3 \right)^{\frac{1}{3}}, \quad (12)$$

其中，每个  $(y) = b - a$  表示持久性  $y = (a, b)$ ，与  $\mu$  的范围是所有的双射从  $F$  到  $G$ 。

为了在定理 2 和 4 更好地理解定理的稳定性结果，我们讨论了两种理论实验。

对于第一个实验中，让  $F_{\lambda} = \{-\lambda, \lambda\}$  和  $G_{\lambda} = \{-\lambda+1, \lambda+1\}$  是两个图一个点并且  $\lambda \in \mathbb{R}_{\geq 0}$ 。这两个点远离对角线随着  $\lambda$ ，同时互相保持相同的欧几

里得距离。因此， $d_{w,p}(F_{\lambda}, G_{\lambda})$  和  $d_{k_0}(F_{\lambda}, G_{\lambda})$  逐渐接近一个常数  $\lambda \rightarrow \infty$ 。相比之下， $d_{kl}(F_{\lambda}, G_{\lambda})$  接近  $\sqrt{\lambda}$  和尤其是无穷的，这意味  $d_{kl}$  强调点在图的高持续性下，作为反映加权每一个  $(u)$  在 (12)。

在第二个实验中，我们比较持久图从两个虚构类 A 的数据样本(即图中  $F, F'$ ) 和 B 的数据样本(即，图  $G$ ) 在。见图 4。我们首先考虑  $d_{kl}(F, F')$ 。正如我们已经在之前的实验中看到， $d_{kl}$  也将因为受高持久性的点影响而发生变化。同样， $d_{kl}(F, G)$  也将受这些点的影响。当  $\lambda$  是足够大的时候。因此，A 和 B 级的实例是密不可分的在最近邻设置。与此相反， $d_B, d_{w,p}$  和  $d_{k_0}$  不强调高持久性点，因此点允许区分类 A 和 B。

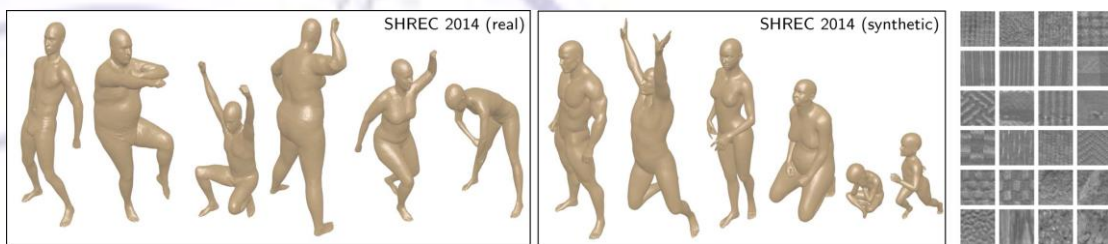


图 3：从 2014 年 SHREC[23] (左, 中) 和 OuTeX Outex\_TC\_00000 [21] (右)。

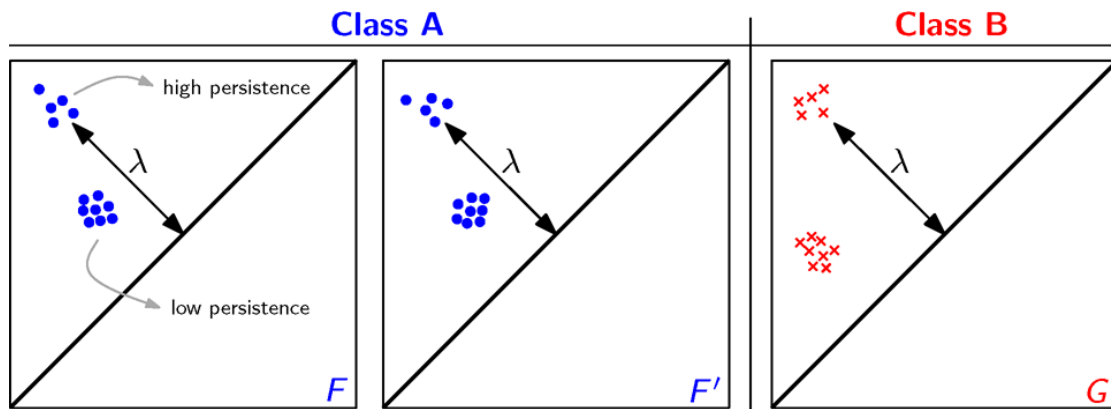


图 4: 两个持续性图  $F$ ,  $F'$ , 从 A 类和一个在 B 类的图  $g$ , 不同点是类只能在它们点的低持久性 (即, 指向靠近对角线)。

## S. 2. 实证结果

我们报告关于两个视觉任务结果, 其中持续同源性已经被证明提供有价值的区别信息 [19]: 形状分类/检索和纹理图像分类。该实验的目的不是要在这些问题上超越最先进的算法思想 - 这将是相当具有挑战性仅仅使用拓扑信息 - 而是说明  $d_0$  和  $d_k$  在  $k^1 d_k$  优势。

**数据集。** 对于形状分类/检索, 我们使用 2014 年 SHREC [23] 标杆, 见图。3. 它由 synthetic 和真实形状 (real shapes), 给出三维网格。该数据的合成部分包含人类的 300 网格 (5 男, 五女, 五个孩子) 有 20 个不同的姿势; 实际包含 40 人的 400 网格 (男, 女) 在 10 个不同的姿势。我们使用全分辨率网格, 即, 没有任何网眼抽取。用于分类, 目标是区分不同的人体模型, 即, 一个 15 级难题题为 SHREC 2014 年 (人造) 和 2014 年 SHREC (真正的) 40 级难题。

对于纹理识别, 我们使用

Outex\_TC\_00000 基准 [21], 降采样到  $32 \times 32$  像素的图像。该基准可提供 100 预定训练/测试分裂, 并且在培训和测试过程中这 24 类相同的使用 10 个图像表示。

**实现。** 对于形状分类/检索, 我们计算经典热内核签名 (Heat Kernel Signature (HKS)) [27] 在一个范围的十次参数增加值  $t_i$ 。对于每个特殊的选择  $t_i$ , 我们在每个对象的表面网格平面上得到一个分段线性函数。正如在讨论秒。3, 我们再计算持久图在规模为 0 和 1 引起过滤。

对于纹理分类, 我们计算 CLBP [15] 描述, (CJ [19])。结果显示 CLBP-Single (CLBP-S) 和 CLBP-Magnitude (CLBP-M) 算子为旋转不变型, 在  $P=8$  的邻域和半径  $R=1$ , 无论算子产生一个标量值响应图像, 可以理解为加权立方体响应图像, 其较低的过滤用于计算持久性图; 见 [28] 的详细信息。

对于这两种类型的输入数据，获得持久性图是使用  $D_{IPHA}$  [2]，它可以直接处理网格和图像。一个标准软间隔支持向量机 C-SVM 分类 [24] 被用在  $L_{IBSVM}$  [5] 实施，这是用于分类。该成本系数  $C$  通过十倍交叉验证训练数据来调整。对于内核  $k_\sigma$ ，它的交叉验证进一步包括内核规模  $\sigma$ 。

### 5.2.1 形状分类

表 1 和表 2 列出  $k_\sigma$  和  $K^L$  的分类结果在 2014 年 SHREC。所有的结果都平均超过 10 个采用随机 70/30 训练/测试分裂的交叉验证运行。具有大致相等的类分布。我们报告的结果只有一维特征；0 维特征导致可比的性能。

在真实和人工合成的数据中，我们观察到， $k_\sigma$  不断改进超过  $K^L$ 。有关  $t_i$  的一些选择，增益甚至范围高达

30%，而在其他情况下，增益是相对较小的。这可以通过输入数据的平滑度看 HKS 时刻  $t_i$  的基本变化来解释。在  $k_\sigma$  的规模  $\sigma$  允许补偿 - 在分类阶段，在一定程度上设置不平滑度。看部分 4。相反， $K^L$  不具备这种能力，并且基本上依赖适当的对输入数据进行预处理。有关  $t_i$  的一些选择，事实上  $K^L$  确实会导致分类精度靠近  $k_\sigma$ 。

但是，当使用  $K^L$  时，我们不得不不要仔细调整 HKS 时间参数，对应于输入数据的变化。这是在大多数情况下不好的，因为 HKS 计算大量网格的顶点的可以相当费时，有时我们甚至可能没有直接使用的网格。改进后的分类率  $k_\sigma$  表明采用附加的自由度实际上对于性能是有益的。

HKS $t_i$	$k^L$	$k_\sigma$	$\Delta$
$t_1$	68.0 ± 3.2	94.7 ± 5.1	+26.7
$t_2$	88.3 ± 3.3	99.3 ± 0.9	+11.0
$t_3$	61.7 ± 3.1	96.3 ± 2.2	+34.7
$t_4$	81.0 ± 6.5	97.3 ± 1.9	+16.3
$t_5$	84.7 ± 1.8	96.3 ± 2.5	+11.7
$t_6$	70.0 ± 7.0	93.7 ± 3.2	+23.7
$t_7$	73.0 ± 9.5	88.0 ± 4.5	+15.0
$t_8$	81.0 ± 3.8	88.3 ± 6.0	+7.3
$t_9$	67.3 ± 7.4	88.0 ± 5.8	+20.7
$t_{10}$	55.3 ± 3.6	91.0 ± 4.0	+35.7

表一：在 SHREC 2014 (人工合成). 的分级性能

HKS $t_i$	$k^L$	$k^G$	$\Delta$
$t_1$	45.2 ± 5.8	48.8 ± 4.9	+3.5
$t_2$	31.0 ± 4.8	46.5 ± 5.3	+15.5
$t_3$	30.0 ± 7.3	37.8 ± 8.2	+7.8
$t_4$	41.2 ± 2.2	50.2 ± 5.4	+9.0
$t_5$	46.2 ± 5.8	62.5 ± 2.0	+16.2
$t_6$	33.2 ± 4.1	58.0 ± 4.0	+24.7
$t_7$	31.0 ± 5.7	<b>62.7</b> ± 4.6	+31.7
$t_8$	<b>51.7</b> ± 2.9	57.5 ± 4.2	+5.8
$t_9$	36.0 ± 5.3	41.2 ± 4.9	+5.2
$t_{10}$	2.8 ± 0.6	27.8 ± 5.8	+25.0

表二：在 SHREC 2014 (实际). 的分级性能

HKS $t_i$	$d_{k^L}$	$d_{k^G}$	$\Delta$	$d_{k^L}$	$d_{k^G}$	$\Delta$
$t_1$	53.3	88.7	+35.4	24.0	23.7	-0.3
$t_2$	<b>91.0</b>	<b>94.7</b>	+3.7	20.5	25.7	+5.2
$t_3$	76.7	91.3	+14.6	16.0	18.5	+2.5
$t_4$	84.3	93.0	+8.7	26.8	33.0	+6.2
$t_5$	85.0	92.3	+7.3	28.0	38.7	+10.7
$t_6$	63.0	77.3	+14.3	28.7	36.8	+8.1
$t_7$	65.0	80.0	+15.0	43.5	52.7	+9.2
$t_8$	73.3	80.7	+7.4	<b>70.0</b>	<b>58.2</b>	-11.8
$t_9$	73.0	83.0	+10.0	45.2	56.7	+11.5
$t_{10}$	51.3	69.3	+18.0	3.5	44.0	+40.5
Top-3 [23]	99.3 - 92.3 - 91.0			68.5 - 59.8 - 58.3		

表三：最近邻检索性能 左：SHREC 2014 (人工合成) 右：SHREC 2014 (实际)

### 5.2.2 形状检索

除了分类实验，我们报告使用标准的评价方法（见[25, 23]）在形状检索性能性能。这允许我们评估内核感应距离  $d_k$  和  $d_{kL}$ 。

为简便起见，仅最近邻性能是在表 3（列出的所有措施的列表，请参阅补充材料）上。用每个形状作为查询用的形状一次，最近邻绩效衡量在检索结果形状属于同一类排名多久一次作为查询。研究调节  $\sigma$  的规模，在柱  $d_k$  栏的最大近邻性能可以取得在一定规模之上。

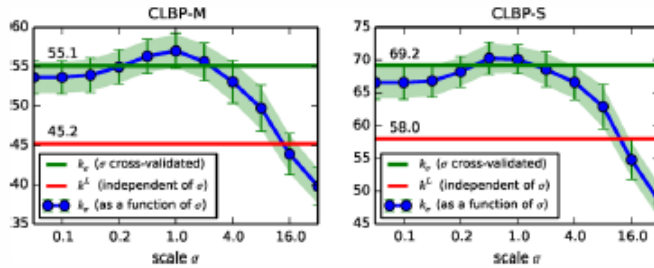
正如我们所看到的，对于分类实验结果是相似的。然而，在几个设定特定的 HKS 时间  $t$ ， $d_{kL}$  执行标准，或者比  $d_k$  好。注意以秒表示。 5. 2. 1

中，这可以解释输入数据的平滑度的改变，由不同 HKS 时间  $t_i$  导致。另一种看法是，近邻性能  $d_{kL}$  围绕 w. r. t 最高结果  $t_i$  相当不稳定。例如，它从 91%（在  $t_2$ ）下降到 53.3% 和 76.7% 在 2014 年 SHREC (人工合成) 和 70%（在  $t_8$ ）下降到 45.2% 和 43.5% 在 2014 年 SHREC (实际)。与此相反， $d_k$  围绕最优  $t_i$  显示稳定的性能。

为了把这些结果用在目前的形状检索工作，表 3 还列出了前三个条目（出自 22）上的相同基准[23]。在真实和人工合成的数据， $d_k$  排在前五的记录。这个表明独立的持续性拓扑是一种丰富的资源在识别信息这个特殊问题。另外，由于我们只评估 HKS 时间参数，性能能够进一步完善通过精心设计的融合策略。

CLBP Operator	$k^L$	$k_{\sigma}$	$\Delta$
CLBP-S	$58.0 \pm 2.3$	<b><math>69.2 \pm 2.7</math></b>	<b>+11.2</b>
CLBP-M	$45.2 \pm 2.5$	<b><math>55.1 \pm 2.5</math></b>	<b>+9.9</b>
CLBP-S (SVM- $\chi^2$ )	$76.1 \pm 2.2$		
CLBP-M (SVM- $\chi^2$ )	$76.7 \pm 1.8$		

表四：在 Out ex\_TC19196H961 的分级性能。



数字 5：纹理分级性能在 SVM 分类器下，随着 (1) 根据  $\sigma$  核  $k$ 。(2) 核  $k$ 。随着  $\sigma$  交叉验证和 (3) 显示核  $k$ 。

### 5.3. 纹理识别

对于纹理识别，在 OuteL\_TC\_00000 1919191919 基准下超过 100 个训练/测试分裂的结果平均。表 4 列出了 SVM 分类器的性能，使用  $k_{\sigma}$  和  $k^L$  在 0 维特征（即，连接部件）。针对此问题，高维特征不是有用的。为了进行比较，表 4 还列出了支持向量机的性能，训练了规范化直方图对 CLBP 的 -S / M 的响应回应，使用  $\alpha \chi^2$  内核。

首先，从表 4，显而易见的是  $k_{\sigma}$  性能更好，比  $k^L$  大幅度的好。在准确性上涨幅高达  $\approx 11\%$ 。第二，很明显的是，对于这个问题，单独拓扑信息对 SVMs 是不具有竞争力的，使用简单的无序算子应对直方图。然而，结果 [19] 表明综合持久性信息（使用持久性景观）与传统的 bag-of-feature 的特征表示导致 state of-

最先进的性能。虽然这表明互补的拓扑特征的互补性，它也表明内核的组合（例如，多内核学习 [13]）可以通过包括核  $k_{\sigma}$  获得更大的收益

为了评估交叉验证（习惯法汇编）的稳定性，选择特定的  $\sigma$ ，图 5 分类性能根据后者。给出的性能曲线的平滑，似乎不太可能通过交叉验证选择参数将会对一个特殊离散化的搜索范围  $[\sigma_{min}, \sigma_{max}]$  产生影响。

最后，我们注意到调整  $k^L$  有相同的缺陷，如在形状分类实验这种情况下，。

虽然在原则上，我们可以使纹理平稳，在 CLBP 响应图像，甚至微调 CLBP 算子，所有的这些战略，需要在开始时的工艺管道进行调整。与此相反，在  $k_{\sigma}$  调整  $\sigma$  T 是在管道的分类器训练结束的时候。

## 六, 结论

我们已经表明, 无论是理论和实证, 即提出的内核对于任务具有的良好作用, 例如采用 SVM 外形分类或纹理识别。此外, 已被证明在实践中调整尺度参数是有益的。

今后工作的一个可能的方向是, 解决为了能够在大规模场景中使用应用程序的计算瓶颈。这可能包括利用可加行和稳定性, 为了使核的价值靠近误差范围内, 特别是, 通过降低内核在总和(10)中的不同点的数量。

而 1-瓦瑟斯坦距离 (Wasserstein distance) 被很好地建立和已经在应用中证明是有益的, 我们希望改善对于持久图  $W. r. t.$  Wasserstein 距离的稳定性。使之超出预期值。这样的结果将会增强我们内核的稳定性, 使用持久性图的基础数据, 这证明了拓扑机器学习全稳定性。

总之, 我们的方法在所有基于内核的机器学习方法中能有效使用拓扑信息。

因此, 这将是有趣的, 看看哪些应用程序领域将因为拓扑机器学习而受益。

## 参考:

- [1] A. Adcock, E. Carlsson, and G. Carlsson. The Ring of Algebraic Functions on Persistence Bar Codes. arXiv, available at <http://arxiv.org/abs/13194.195319>, 2013.
- [2] U. Bauer, M. Kerber, and J. Reininghaus. Distributed computation of persistent homology. In ALENEX, 2014.
- [3] P. Bubenik. Statistical topological data analysis using persistence landscapes. arXiv, available at <http://arxiv.org/abs/12197.6437>, 2012.
- [4] G. Carlsson. Topology and data. Bull. Amer. Math. Soc., 46:255-308, 2009.
- [5] c.-c. Chang and c.-J. Lin. UBSYM: A library for support vector machines. ACM TIST, 2(3):1-27, 2011.

- [6] F. Chazal, L. Guibas, S. Oudot, and P. Skraba. Persistence-based clustering in Riemannian manifolds. In SoSC, 2011.
- [7] C. Chen, D. Freedman, and C. Lampert. Enforcing topological constraints in random field image segmentation. In CVPR, 2013.
- [8] M. Chung, P. Bubenik, and P. Kim. Persistence diagrams of cortical surface data. In IPMI, 2009.
- [9] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Compo Ceom.*, 37(1): 103-120, 2007.
- [10] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko. Lipschitz functions have  $L_p$ -stable persistence. *Found. Comput. Math.*, 10(2):127-139, 2010.
- [11] H. Edelsbrunner and J. Harer. *Computational Topology. An Introduction.* AMS, 2010.
- [12] M. Gao, C. Chen, S. Zhang, Z. Qian, D. Metaxas, and L. Axel. Segmenting the papillary muscles and the trabeculae from high resolution cardiac CT through restoration of topological handles. In IPMI, 2013.
- [13] M. Gonen and E. Alpaydin. Multiple kernel learning algorithms. 1. *Mach. Learn. Res.*, 12:2211-2268, 2011.
- [14] L. Greengard and J. Strain. The fast Gauss transform. *SIAM J. Sci. Statist. Comput.*, 12(1):79-94, 1991.
- [15] Z. Guo, L. Zhang, and D. Zhang. A completed modeling of local binary pattern operator for texture classification. *IEEE TIP*, 19(6):1657-1663, 2010.
- [16] T. Iijima. Basic theory on normalization of a pattern (in case of typical one-dimensional pattern). *Bulletin of Electrical Laboratory*, 26:368-388, 1962.
- [17] R. J. j. Iorio and Y. de Magalhaes Iorio. *Fourier analysis and partial differential equations.* Cambridge Stud. Adv. Math., 2001.
- [18] A. Krizhevsky, I. Sutskever, and G. Hinton. ImageNet classification



with deep convolutional neural networks. In NIPS, 2012.

[19] C. Li, M. Ovsjanikov, and F. Chazai. Persistence-based structural recognition. In CVPR, 2014.

[20] D. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91-110, 2004.

[21] T. Ojala, T. Maenpaa, M. Pietikilinen, J. Yiertola, I. Kyllonen, and S. Huovinen. OuTeX - new framework for empirical evaluation of texture analysis algorithms. In ICPR, 2002.

[22] D. Pachauri, C. Hinrichs, M. Chung, S. Johnson, and Y. Singh. Topology-based kernels with application to inference problems in Alzheimers disease. IEEE TMI, 30(10):1760-1770, 2011.

[23] Pickup, D. et al. SHREC '14 track: Shape retrieval of non-rigid 3d human models. In Proceedings of the 7th Eurographics workshop on 3D Object Retrieval, EG 3DOR' 14. Eurographics Association, 2014.

[24] B. Scholkopf and A. J. Smola. Learning with Kernels: Support

Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge, MA, USA, 2001.

[25] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The Princeton shape benchmark. In Shape Modeling International, 2004.

[26] P. Skraba, M. Ovsjanikov, F. Chazal, and L. Guibas. Persistencebased segmentation of deformable shapes. In CVPR Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment, 2010.

[27] J. Sun, M. Ovsjanikov, and L. Guibas. A concise and probably informative multi-scale signature based on heat diffusion. In SCP, 2009.

[28] H. Wagner, C. Chen, and E. YU9ini. Efficient computation of persistent homology for cubical data. In Topological Methods in Data Analysis and Visualization II, Mathematics and Visualization, pages 91-106. Springer Berlin Heidelberg, 2012.