

指导教师： 杨涛

提交时间： 13/3/ 21

# CVPR2015 Paper Translation

No: 01

姓名： 巩固

学号： 2013302520

班号： 10011303



# 一种新的局部线性模型的视觉识别方法

Qingfeng Liu, Chengjun Liu

New Jersey Institute of Technology Newark, NJ, USA ql69, cliu@njit.edu

## 摘要:

本文提出了一种新的局部线性法模型,有了目标,不仅发展高效表示和分类方法,还建立关系,它们之间以近似某种分类规则,如贝叶斯决策规则。走向那末。首先,该模型表示的测试样本作为一个线性,所有的训练样本的组合和派生学习系数的新表示重建,局部性和稀疏性的限制。这个理论分析表明,新的表示最近邻的分组效应,这是能够近似“理想表示”。然后在本地线性分类器 KNN 模型 (11knc),这显示其连接到最小的贝叶斯决策规则在核中的错误,提出分类。此外,局部线性最近均值分类器 (11nmc),其关系到 11knc 是就像最近均值分类器 KNN 分类器,是衍生。此外,提供可靠的内核密度估计,变功率变换和系数截止方法应用于提高性能所提出的方法。建议的有效性几种视觉识别任务模型进行评价如人脸识别,场景识别,物体识别动作识别。实验结果表明该模型是有效的,优于一些其他有代表性的流行方法。

## 1. 介绍

视觉识别,如人脸识别,对象识别,场景识别和动作识别在过去几十年中受到广泛关注。许多方法开发和最成功的一个好的研究方法是子空间方法,无论是线性子空间[ 2 ], [ 27 ] 或非线性[ 16 ], [ 35 ]。最近,稀

疏表示为基础的方法[ 31 ]是提出解决问题的鲁棒表示分类。许多变种 [ 11 ], [ 34 ], [ 37 ]建议将歧视性的信息学习字典和稀疏表示。对于大多数的方法,表示和分类都是独立开发的,这违反了需要表示方法应提供服务 and 方便视觉识别的后续分类方法。此外,这些方法带来的其他问题,如分类限制,计算复杂度等。为了解决这些问题,我们提出了一个新的本地线性 k 近邻 (11knn) 模型的鲁棒性视觉识别。所提出的方法首先学习一个新的测试样本的线性组合表示基于重构准则的所有训练样本,局部性和稀疏性。新的表示向量,具有最近的分组效应邻居,然后提供作为输入的局部线性 KNN 模型为基础的分类 (11knc) 和局部线性最近的平均分类 (11nmc),分别。力量该 11knc 保证是建立最小误差的贝叶斯决策规则核密度估计的观点。移动电源变换和系数截断法用于鲁棒性和可靠性。所提出的方法的有效性进行评估五代表数据集。特别是人脸识别,人脸数据库[ 17 ]用于场景识别,场景数据集[ 21 ]应用;物体识别,加州理工学院 256 个数据集[ 10 ]利用;和行动识别,的 ucf50 集[ 22 ]是用。实验结果表明所提出的方法的可行性。该系统的体系结构如图 1 所示。这个花纹矢量进行预处理的移动电源改造。然后降维方法应用。该 11knn 模型进一步推导出一个新的表示向量 V 最后的 11knc 和 11nmc 用于分类。

本文的主要贡献如下  
首先,我们提出了一种新的局部线性

KNN 1。所提出的局部线性 KNN 模型的体系结构

(11knn) 模型, 导出表示可以近似地利用“理想表示”所提出的模型。

•其次, 我们提出了局部线性法模型分类器 (11knn), 揭示其连接最小误差的贝叶斯决策规则内核密度估计的背景。

第三, 我们提出了局部线性最近的平均分类 (11knn), 其关系到 11knn 只是类似最近的平均分类之间的关系和 KNN 分类器。

第四, 我们解决了可靠的密度估计问题应用转向功率变换和系数截断法: 对全球带宽的敏感性和不良遥远的邻居的影响。

## 2. 相关工作

稀疏表示方法被广泛应用于视觉识别。有些方法 [ 5 ], [ 6 ] 寻求模型字典内的类内变化以提高人脸识别的性能。最近, 一些歧视字典学习方法提出了稀疏表示。张等。 [ 37 ] 提出一个目标函数并应用了一种判别奇异值分解 (d-ksvd) 学习的判别方法字典和分类同时。江等

铝。 [ 11 ] 通过增加一个标签的一致性提高了 37 正则化项。周等。 [ 38 ] 提出一个联合

字典学习 (JDL) 的方法, 共同学习一个共同的字典和类特定 subdictionaries 加强对字典的歧视。杨等人。 [ 33 ], [ 34 ] 提出了渔民歧视字典学习 (网) 的方法学由一组特定的类构成的字典子字典。在比较中, 我们的方法提高了这些方法在下面的方法。(一) 字典的推导

其他的方法是非常耗时的, 因为它需要反复更新字典和稀疏表示或者。

(II) 有些方法是受限制的线性分类, 不包括非线性分类可以取得更好的性能。(第三) 本次字典基于方法可能会导致性能恶化时每个类的训练样本的数目是小的因为这些子字典是单独训练的每个类或依赖于相应的类太多。

(IV) 我们所提出的方法能够建立这种关系

表示方法及其分类在逼近贝叶斯决策规则的意义最小误差。

## 3. KNN 模型的局部线性

在文献中的一个共同的假设是, 一个测试样本是所有训练样本的线性组合 [ 31 ] 捕捉到真实数据集的变化。理想的情况是, 只有训练的系数样品与同一类标签为测试样本非零, 否则为零。数学上, 思想表示被定义如下。

定义 3.1 理想的表示。给出了测试样品的  $X \in \mathbb{R}^M$  来自班级的, 所有的训练样本  $x_i \in \mathbb{R}^M$

( $i = 1, 2, \dots, M$ ), 理想的代表性是系数向量  $v = [v_1, v_2, \dots, v_M] \in \mathbb{R}^M$  我是如此  $x = \sum_{i=1}^M v_i x_i$  (1) 第六号非零, 如果  $x$  属于  $C$ -th 类及其他 0。

其结果是, 理想的表示是高度稀疏的, 这诱导发展的稀疏表示为基础的方法 [ 31 ]。

然而, 以下 2 个固有的问题基于稀疏表示的方法仍在等待一个溶液。首先, 稀疏约束不能保证预期系数是非零的近似理想表示。具体而言, 训练样本这是在同一类的测试样本经常

高度相关, 而稀疏约束往往倾向于选择一个非零系数 [ 39 ]。二, 基于稀疏表示的分类方法不直接相关的最佳分类规则, 如贝叶斯最小误差的决策规则 [ 7 ]。为了解决这些问题, 我们提出了一个新的本地线性 KNN (11knn) 基于观察模型最近训练样本邻居的测

试样本

极有可能分享同一个类标签测试样本，如果他们是稳健选择。注意我们方法采用鲁棒的 L1 范数和利用最近训练样本邻居的测试样本的地方财产。

结果，我们的方法揭示了分组最邻近效应（健），因此，理想的代表性更容易被我们的提出方法。

在数学上，小说的局部线性 KNN (11knn)

模型定义如下：

$$\min_v \left\| X - BV \right\|_2 + \lambda \left\| V \right\|_1 + \alpha \left\| V - \beta D \right\|_1 \quad (2)$$

其中  $X \in \mathbb{R}^R$  在训练样本矩阵和系数向量  $V \in \mathbb{R}^m$  是派生的表示。 $\left\| \cdot \right\|_2$  是 L2 范数和  $\left\| \cdot \right\|_1$  是 L1 范数。向量  $a = [D_1, D_2, \dots, \text{糖尿病}]^T \in \mathbb{R}^m$ ，和  $D = \left\{ \left\| X - X_i \right\|_2 \right\}$ 。参数  $\sigma$

用于调节衰减速度。我们可以观察到是较大的，如果训练样本双接近试验样品

X.

第一项保持重建能力，二是保持鲁棒性的稀疏约束属性，和第三个术语表示的地方属性使训练样本越接近测试样品，其系数越大。模型参数： $\lambda$ ， $\alpha$  和  $\beta$  有助于平衡每个学期。这

建议的模型，从而强调最近的训练样本测试样本的邻居，并分配给它们大的系数。第 3.3 节进一步表明，只有最大的在同一个最近的训练邻居的系数作为测试样本的类是实现良好的

性能。所提出的方法的基本原理是双重的。首先，最近的训练样本邻居更多可能是在同一类的测试样本，如果他们强劲选择。此外，如图 3.1 所示，该模型具有分组效应最近邻居（健）的训练样本，这是高度相关的，接近测试样品往往具有相似的大系数。作为结果，该模型往往得出的理想表示。

• 二的 11knn 模型为基础的分类 (11knn) 近似的贝叶斯分类器核密度估计在 3.2 节中的意义基于所提出的表示模型。

### 3.1. 表示

新的表示是通过优化的标准方程 2. 的 FISTA (快速迭代收缩阈值算法) 算法 [1] 适用于求解它。方程 2 可以分解为五 (五) + 克 (五)，其中  $f(v) = \left\| X - BV \right\|_2 + \alpha \left\| V - \beta D \right\|_1$  和  $G(V) = \lambda \left\| V \right\|_1$  为了保证收敛，为 FISTA 最大步长算法被选为 1L，其中  $L = 2 \lambda \text{Max}(\text{肺结核} + I)$ ，这意味着矩阵的最大特征值的两倍肺结核 + 一个

通过优化方程 2 中的标准，新的表示具有最近邻的分组效应（健）如下图所示。

定理 3.1 给出了一个二级标准化测试样本 ( $\left\| X \right\|_2 = 1$ )，L2 归一化后的训练样本矩阵

$B = [B_1, B_2, \dots, B_m]$  和向量  $a = [D_1, D_2, \dots, \text{糖尿病}]^T$ ，让我们来  $v = [V_1, V_2, V_3, \dots, V_m]^T$  不是解决方案

在方程 2 中定义的 11knn 模型。定义样本相关  $\rho$  两  $B_i$  和  $B_j$  作为训练样  $\rho = \text{Cov}(B_i, B_j) / (\text{Cov}(B_i, B_i) \text{Cov}(B_j, B_j))$  和系数  $V$  之间的差异  $v_i - v_j$  (我, 2, 1, ..., M (i, j) =  $\left\| v_i - v_j \right\|_2$ )

我 -  $v_j$  (3) 然后，如果是伏的迹象  $v_i - v_j$  是一样的，我们有  $M(i, j) \leq \text{Cov}(B_i, B_j) + \beta \left\| D_i - D_j \right\|_2$  (4 在哪里  $P = (1 + 2 \left\| \alpha \beta \right\| D) / 2$ )，这是一个常数方程 4 被称为最近的分组应邻居 (生殖)，这意味着，如果训练样本高度相关 ( $\rho \approx 1$ ) 和接近试样品 ( $D_i \approx D_j$  和二,  $D_j$  大)，那么系数的训练

样本是相似的 ( $\sum_{j=1}^M V_j \approx \sum_{j=1}^M V_j$ )。实验在 5.5 节的分析, 意味着密封性束缚在方程 4, 生殖性放大训练样本在同一类中的系数测试样本, 同时抑制其他。这是案例与理想的代表性相一致。

### 3.2. 分类

在我们得到的  $11knn$  表示  $V$  给定训练样本的测试样本的模型矩阵  $B$ , 我们首先提出了基于的  $11knn$  模型—费里 ( $11knc$ ) 分类测试样本  $X$  的分类—阳离子规则定义如下:  $C^* = \arg \max_{C \in BC} CX$  第六 (5) 一千三百三十一在哪里  $= 1, 2, \dots$ , 是班级的标签, 是前一组训练在  $C$ -th 类样本。这样的  $11knc$  分配测试样本的类\*由所有主要软投票在每个类的训练样本, 这意味着测试样本被分类为拥有最多的总和的类系数。有效的  $11knc$  表示定理

2.2。用一些合理的近似, 分类方程 5 中定义的规则近似贝叶斯决策最小误差准则 [ 7 ]。

定理 3.2 给出了测试样本, 相应的表示五, 转换为  $6 = 6 - V_{min} - V_{min} V_{max}$  和  $6 = 我 = 1$

首先被应用, 其中  $V_{min}$  和  $V_{max}$  最小和最大值之间的所有元素向量  $V$

然后, 如果事先分配的磷 (碳) 是相等的所有类, 贝叶斯决策规则的建议在内核上  $11knn$  模型为基础的分类密度估计。  $C^* = \arg \max_{C \in BCVI} CX \approx \arg \max_{C \in BC} CX \beta$  迪+常  $\propto \arg \max$

$CP (C | X)$  (6) 请注意转换  $\sum_{C \in BC} 六$  秋天  $0, 1$ ] 为了建立与后验概率的关系。这是很容易看到的转换不影响  $11knc$  的结果。另一种分类方法, 局部线性最近均值分类器 ( $11nmc$ ), 以及建议。关系  $11nmc$  和  $11knc$  之间类似于之间最近的平均分类和  $KNN$  分类器。我们首先定义“平均”的  $C$ -th 类如下:

主持人  $= X \text{ 双} \in BC \vee IBI$  (7) 的  $11nmc$  然后定义如下  $C^* = \arg \min_{C \in BC} | | X - MC | |$   
 $2 \text{ 二} = \arg \min_{C \in BC \vee IBI} | | X - X \text{ 双} \in BC \vee IBI | |$   
 $2 \text{ 二}$

(8)

请注意, 流行的最小残留分类 [ 31 ] 是一个特殊的情况下, 我们提出的  $11nmc$  当地方

信息被丢弃。我们提出的区别  $11nmc$  和最小剩余的分类: 使用新的  $11nmc$  派生的表示而不是

最小残差分类器的稀疏表示。

### 3.3. 可靠的内核密度估计

定理 3.2 告诉我们,  $11knc$  动力来源从内核密度估计。然后有 2 个问题核密度估计应解决的问题为了提高系统的可靠性, 即对全局窗口宽度表示为  $\sigma$  价值和遥远的邻居的影响。

第一个问题是核密度估计当基础密度时, 受到全局窗口宽度需要不同数量的平滑在不同的位置, 这意味着  $\sigma$  值应该是不同不同位置的特点。正如在 [ 28 ], 内核中所证明的密度估计的作品以及密度不远从高斯的形状, 因为均匀使用全局窗口宽度。转换的功率变换能够将数据转换为近高斯形状, 使新的数据可以很好地估计。因此, 我们申请转移电源转换为模式向量

在应用  $11knn$  模型。  $T(x) = X + \lambda 1E$  |  $\lambda 2$  符号  $(x + \lambda 1E)$  (9 符号  $(\times)$  表示每个元素的符号向量向量  $x$  的值为  $0, 1$  和  $1, 0 < \lambda 1, \lambda 2$  和电子  $= [ 1, 1, \dots, 1 ]$  吨。请注意所有的向量运算是元素明智。此外, 在新的转变空间, 进一步减轻敏感性参数  $\sigma$ , 我们也提出申请  $L2$  归一化变换空间模型中的向量空间。作为结果, 我们发现  $\sigma$  值没有影在实践中的表现(见第 5 节)。

另一个问题是, 在新的转换空间, 我们可以丢弃一些遥远的邻居, 其中有尾随系数可能对性能有不利影响。因

此，我们提出了一个系数截断法，在那里只有最顶级的第六每个类的值保持这样分类可以更有效地计算和密度估计更可靠。这样的  $l_1$ knnc 定义如下

$$C^* = \arg \max_C \sum_{(i \in BC) \wedge (i \in t(k))} C_i$$

第六 (10)

在何处 (克) 是一组顶部的最大值为每一个类。同样，在  $l_1$ nmc 定义如下：

$$C^* = \arg \min_C \sum_{(i \in BC) \wedge (i \in t(k))} |X_i - X_i| + \lambda |C_i|$$

二

(11) 在实践中，对绩效的重要价值 (见 5 节)。

图 2。数据集的样本图像：(一) 应为人脸数据库，(b) 15 个场景数据集，(C) 的 mit-67 室内场景数据集，(D) 加州理工学院的 256 个数据集和数据集的 ucf50 (E)。

## 4. 实验

在这一节中，我们评估的性能提出的  $l_1$ knn 模型和两个分类器的  $l_1$ knnc 和多种视觉识别数据库的  $l_1$ nmc：脸人脸数据库中的人脸识别 17 在 15 个场景数据集 [13] 和 mit-67 室内场景数据集 [21]，加州理工学院 256 数据集 [10]；在 ucf50 集动作识别 [22]。一些样本图像如图 2 所示。图像或视频首先被表示为一个模式向量。为了进行公平比较或达到可比性结果国家的最先进的方法，我们使用不同的不同数据集的模式向量。请查看更多细节在相应的子部分。和边缘 Fisher 分析 (MFA) 与主成分分析 (主成分) 是用来减少尺寸和提取功能。

## 4. 1. 人脸识别

我们评估所提出的人脸识别方法通过使用人脸数据库，它是由 126 个正面的图片为 4000 个人，每 26 个在两会上拍摄的照片。数据的一个子集 [17]，其中包括 50 名男性和 50 名女性的图像尺寸  $165 * 120$ ，是从原始数据库中选择的。然后，我们遵循在 [31] 中定义的实验设置和 [34] 作出公平的比较，14 幅图像照明变化和表达式为每个选择人：七张 1 张照片，用于培训和其他七从会话 2 进行测试。应用前所提出的方法，人脸矢量的维数是减少到 180。

模型参数作为  $\sigma = 1$ ， $\lambda = 0.020.1$ ，和 1.5 的建议模型。为实验设置 1 精度  $d$ -ksvd [37] 85.40LC-KSVD [11] 89.7JDL [38] 91.7 其中 [34] 92SRC [31] 94.99 该  $l_1$ nmc 96.14 该  $l_1$ knnc 97

表 1。该  $l_1$ knnc 之间的比较， $l_1$ nmc 和其他流行的方法在面对数据库。方法准确率  $81.40 \pm KSPM$  [13] 0.50 然后 [32]  $80.28 \pm 0.93$  有限责任公司 [30]  $80.57 \pm -KC$  [9]  $76.67 \pm 0.93$

$d$ -ksvd [37] 89.10LC-KSVD [11] 90.40laplaciansc [8] 89.7 该  $l_1$ nmc  $97.45 \pm 0.27$  该  $l_1$ knnc  $93.54 \pm 0.45$

表 2。该  $l_1$ knnc 之间的比较， $l_1$ nmc 和其他流行的方法在 15 个场景的数据集移动电源转换， $\lambda_1 = 0$ ， $\lambda_2 = 0.9 \lambda$ 。对的  $l_1$ knnc， $k = 5$  的值为  $l_1$ nmc，的值 = 7。在表 1 中给出的结果表明，所提出的方法是能够提高后其他流行的方法显著

## 4. 2. Scene Recognition

### 4. 2. 1 The 15 Scenes Dataset

15 个场景数据集 [ 13 ] 包含 4485 个图像从 15 个场景类别, 每一个图像的数量范围从 200 到 400。实验协议定义在 [ 13 ] 和 [ 32 ] 中, 我们随机选择 100 幅像每类为训练和剩余的 10 次迭代测试。首先, 我们使用的空间传销表示所提供的 [ 11 ] 以表示图像为载体 3000 维度。表示是通过使用一四级空间锥体和一个码本的大小 200。然后, 图像矢量进一步降低到维 500。对于移动电源转换,  $\lambda_1 = \lambda_2 = 0.5$ 。模型参数作为  $\lambda = 0.050.1$ , 和 1。为 11knc, 价值  $k = 1$  的 11nmc,  $k = 2$  的值。它可以从表 2 的结果, 我们提出的结论方法是能够取得更好的结果比非线性或基于线性核的支持向量机所用的比较方法。

表 3

表 3。该 11knc 之间的比较, 11nmc 在 mit-67 室内场景的其他流行的方法数据集

#### 4.2.2 The MIT-67 indoor scenes Dataset

mit-67 室内场景数据集 [ 21 ] 是一个很大的挑战场景识别数据集, 其中包含 67 个室内类有 15620 张图片。我们遵循常用的实验设置 [ 21 ], 其中  $80 * 67$  图像被使用用于训练和  $20 * 67$  图像的测试。性能测量的平均分类准确率在所有类别。我们考虑的是, 鱼矢量特征 [ 26 ] 表示。第一次投影的筛选功能 256 个维度和 80 个视觉的码本语言是计算的, 然后是二维的是  $2 * 256 * 80 = 40960$ 。然后, 我们进一步降低尺寸到 2000。对于转向功率变换,  $\lambda_1 = 0, \lambda_2 = 0.5 \lambda$ 。选择模型参作为  $\lambda = 0.01, \alpha = 0.1$ , 和  $\beta = 1.5$  时的  $11\beta = 0.5$  为最佳性能的 11knc。对无论是 11knc 和 11nmc,  $k = 20$  的

值。表 3 中的结果表明, 该方法是能够

我们能够实现结果 mit-67 极具竞争力的挑战室内场景数据集。请注意, 我们学习的渔民直接从图像的特征, 而不是学习的一部分探测器 [ 12 ]。此外, 我们减少尺寸 40960 至 2000 的, 这节省了大量的存储空间。并没有使用数据增强技术。不过, 我们仍然可以取得非常有竞争力的结果国家的最先进的方法 [ 12 ]。

#### 4.3. Object Recognition

加州理工学院的 256 个数据集 [ 10 ] 包含 30607 个图像分割分成 256 个对象类和一个杂类。我们遵循共同的实验设置 [ 30 ], 30, 45, 15, 60 个图像, 每类随机选择训练并没有超过 25 个图像的测试 3 次迭代

表 4

表 4。所提出的方法和其他的比较流行的方法在加州理工学院 256 数据集。

为了达到国家的最先进的方法, 所提出的方法是建立在 4096 个维度利用预训练卷积提取的向量神经网络 cnn-m [ 4 ]。移动电源转型,  $\lambda_1 = 0, \lambda_2 = 0.5 \lambda$ 。我们进一步减尺寸为 1000。选择模型参数作为  $\sigma = 1.5, \lambda = 0.01, \alpha = 0.1$  和  $\beta = 1.5$ 。两的 11knc 和 11nmc,  $k = 15$  的值。这个结果显示在表 4 中, 演示了建议方法是能够与其他流行的方用支持向量机进行分类训练图像大小。请注意, 所提出的方法不使用数据增强和微调技术 [ 4 ]。然而我们仍然可以取得比较的结果国家的最先进的方法 [ 36 ], [ 4 ]。

#### 4.4. Action Recognition

我们用行动识别数据集的数据集: ucf50 [ 22 ], 这是一个大规模的视频

数据集的动作识别收集来自 YouTube, 评估建议方法。它由 50 个动作类别组成 6676 个视频和至少 100 个视频为每个动作类。我们遵循的实验设置在 [ 23 ] 中, 将数据集划分为 5 组相似尺寸数据, 采用分组交叉验证法。请注意, 此设置比 LEAveone 交叉验证更多的挑战在 [ 22 ] 提出的 25 倍。该行行动银行功能 [ 23 ] 适用于代表视频数据与尺寸 14965, 进一步减少 500。对于移动电源转换,  $\lambda_1 = 0.0$   $\lambda_2 = 0.8$ 。模型参数作为  $\sigma = 1.5 \lambda = 0.05$ ,  $\alpha = 0.1$ , 和  $\beta = 1.5$ 。无论是 11k 和 11nmc,  $k = 10$  的值。表中的结果 5 表明, 所提出的方法可以提高其他流行的方法很多。

## 5. 分析

在这一节中, 我们提供了更全面的分析关于性能的建议方法。

表 5

特别是, 我们评估所提出的方法在以下关键问题: (1) 转移的有效性功率变换, (2) 敏感性参数  $\sigma$ , (3) 参数  $K$  的敏感性 (4) 比较普通的 KNN 分类器, 和 (5) 分组评价最近邻效应。和所有的实验在本节中的设置与所用的相同除另有规定外, 上述试验段。

### 5.1. The effectiveness of the shifted power transformation

首先, 我们评估的有效性的转移功率变换 (SPT) 通过比较结果

“11knnc”(已经与 SPT), “没有 11knnc SPT”, “11nmc”(已经与 SPT) 和 “11nmc 没有 SPT” 在不同的值的参数, 这是定义在 3.3 节。没有规范, 当我们表示方

法为 “11knnc” 或 “11nmc” 移动电源改造已应用。如图 3 所示使用移动电源的性能确实得到改善在所有的价值观下的转变参数是固定的。

Figure 4

### 5.2. The sensitiveness to the parameter $\sigma$

如第 3.3 节所提出的, 转向功率变换二语规范化能缓解  $\sigma$  敏感的参数, 它代核密度估计的全局窗口宽度。如图 4 所示, 我们展示了结果 “11knnc” 和 “11nmc” 的不同价值观下参数  $\sigma$ 。参数  $\sigma$  真正不影响性的分类, 因为我们已经申请了这两个转换后的功率变换与二语规范化到矢量  $D$ , 但是, 没有移动电源改造, 性能依赖于  $\sigma$  价许多。我们也发现, 二语规范化的向量维是必要的, 以保证性能, 否则性能将下降到 10% 以下。进一步表明, 敏感性参数  $\sigma$  我们所提出的方法可以缓解, 一些极端对  $\sigma$  如 20 值, 30 进行了评价。性对 11knnc 58.40 (比这更好的报道以上) 和 11nmc 性能均为 58.49 价值观。

### 5.3. The sensitiveness to the parameter $k$

如第 3.3 节, 系数截断方法是能够丢弃较长距离的邻居其中有助于有不利的尾随系数对性能的影响。因此, 我们评估的性能当参数的值的变化和所有的其他参数是固定的。结果如图 3 所示表明我们不能的价值的的重要性使用太小的价值, 也没有必要使用更大的价值

### 5.4. Comparison to KNN Classifier



我们比较我们所提出的分类器 KNN (K=3) 分类作为训练数据的量在这变化剖面。表 6 的结果表明，我们提出的分类器可以提高在平原 KNN 分类器显著。

表 6

### 5.5. The Grouping Effect of the Nearest Neighbors

在这一节中，我们评估了绑定的松散度最近邻的分组效应与目标表明该模型能够得到 11knn 接近理想表示的表示在测量方面：真正的激活率（焦油）和假激活率（远）。给定的测试样本的测试数第  $i$  个样本预期的非零系数  $T_i$ ，第一个测量：真正的活化比（焦油），被定义如下

$$p_{ntest} = \frac{1}{N} \sum T_i$$

(12) 这意味着预期系数平均数—客户被激活（非零）的所有测试样本。理想的情况下，焦油是训练样本的大小  $C$ -th 类如果测试样品是从  $C$  类二是假激活率（远）如果是  $p_{ntest} = \frac{1}{C-1}$  测试

(13) 在哪里是类和网络连接的数量是多少非预期的第  $i$  个测试样本的非零系数。远代表的期望平均数被激活的所有其他的系数伪类 ( $C-1$ ) 所有试验样品。我们期待提出的模型保持较高的值焦油，是接近在每个类和一个较低的训练样大小价值远，接近 0。如表 7 所示，我们提出的准确性，价值远和焦油的值时的参数  $\lambda$ ， $\alpha$  和变化。可以观察到，有时增加（减少）焦油的值也会导致增加（减少）的价值，可能会降低性能。因此一个权衡之间的价值焦油的价值是必要的，以达到最佳性能

### 6. 结论

本文提出了一种新的局部线性 KNN 模型鲁棒视觉识别，

表 7

理论分析表明派生的表示具有分组效应最近的邻居，这是能够近似的“理想表示”。然后局部线性算法的基础分类器 (11knn)，并证明了近似核密度估计的贝叶斯分类器，提出分类。此外，局部线性最近的平均分类 (11nnc)，提出了。此外，移动电源转换和系数—所有的切断方法是用来改善性能建议的分类。建议的有效性几种具有代表性的视觉模型进行评价识别数据库，实验结果显示该模型优于其他一些代表流行方法。

### References

1) 和贝克和米。teboulle。iterative 和快速 shrinkagethresholding linear inverse 算法的问题。暹罗 J. 成像科学, 2 (1): 183 - 202, 2009.

2) belhumeur, P. N. J. P. hespanha 克里格曼和 D. J. 一系列的特征脸的识别与 fisherfaces: 使用特定的类。linear 投影。IEEE 跨。模式分析。马赫数。intell., 19 (7): 711 - 720, 1997.

(3), 升任。博, X, 和 d. 福克斯。multipath sparse 编码。hierarchical 语言学的追求。在对苏铁 2013 IEEE 计算机视觉和模式识别的会议, 作者 660 - 667, 2013 年。

[ 4 ]。 chatfield, simonyan vedaldi, zisserman. 返回的魔鬼在细节: delving convolutional 深。网。在英国机器视觉对苏铁会议 (2014 bmvc)。6

- [ 5 ] W. 邓, 胡, J 和 J. 郭。SRC: undersampled 扩展通过 intraclass 脸识别字典的变种。IEEE 反。模式分析。马赫数。intell., 34 (9): 1864 - 1870, 2012。2
- [ 6 ] W. 邓, 胡, J 和 J. 郭。在国防部 sparsity 为基础的脸识别。在 cvpr 页, 399 - 406, 2013 年。2
- [ 7 ] 关于哑弹, R. P. D. E. 哈特和 G. 鹤。模式分类 (第二版)。威利-跨学科, 2000。3, 4 1336
- [ 8 ]。高。W. H. 曾荫权, 属不嘉。laplacian sparse hypergraph laplacian sparse 编码, 编码, 和应用。IEEE 跨。模式分析。马赫数。intell., 35 (1): 92 - 104, 2013 年。5
- [ 9 ] (J. C. gemert, J. M. geusebroek, C. J. veenman, 和 W. smeulders)。codebooks categorization 内核的场景。在 eccv, 696 - 709 页, 2008。5
- [ 10 ] 克。格里芬, 一只鸽子, 庇隆和 P。。。加州理工学院 256 对象类别数据集。7694 加利福尼亚技术研究所的报告, 科技, 2007。1、5、6
- [ 11 ]。江林, Z 和 L. 戴维斯。标签: consistent - 坡 discriminative 词典的学习和识别。IEEE 反。模式分析。马赫数。intell., 35 (11): 2651 - 2664, 2013 年。1, 2, 5, 7
- [ 12 ] juneja, vedaldi M. C。在 jawahar, 和 zisserman。。。地铁的呼喊: 独特的部分的场景分类。在 2013 年 IEEE 计算机视觉和模式会议
- 议 识别, 波特兰, 或美国, 六月 23 - 28, 2013 年, 页 923 - 930, 2013 年。6
- [ 13 ]。施密德, Lazebnik, C 和 J. 庞塞。超越的。语言: 英语语言学 recognizing 自然空间金字塔 场景类别。在 cvpr, 2169 - 2178 页, 2006。1, 5
- 如果 [ 14 ] 属苏 J, E, H. P. 兴, 如果 F 和 F. 银行和 highlevel 对象: 表示的图像场景分类和 semantic sparsification 特征。在 夹页, 1378 - 1386, 2010。6
- [ 15 ] 问: 如果 J. 吴, 和从中收获的水平。这里的视觉概念。从大规模的互联网图像。在 IEEE 会议 2013 他是计算机视觉和模式识别, 波特兰, 或, 美国, 六月 23 - 28 页, 851 - 858, 2013 年。6
- [ 16 ] C. 刘。他 dimensionality capitalize increasing 技术 为提高大挑战的面的识别性能。IEEE 跨。模式分析。马赫数。intell., 28 (5): 725 - 737, 2006。1
- [ 17 ]。mart'mez 和 M. C. 的卡卡。PCA 和 LDA。IEEE 跨。模式分析。马赫数。intell., 23 (2): 228 - 233, 2001。1, 5
- [ 18 ] Piedra 橄榄和和。。。场景建模 的形状和位置: holistic 表示空间的信封。国际 计算机视觉的杂志, 42(3): 145 - 175, 2001。7
- 潘迪 [ 19 ] 和 Lazebnik 与 M。。。场景 识别和 weakly 监督对象的一部分为基础的本地化的 deformable 模型。他在 IEEE 计算机视觉国际会议,

- iccv 2011 巴塞罗那, 西班牙, 6 月 13, 2011  
1307 - 1314 页, 2011. 6
- [ 20 ] F. 桑切斯和 perronnin, J. T. mensink. 写的。”  
费舍尔内核的大规模的图像分类。在对  
第十一式: 欧洲计算机视觉会议  
第四部分, 作者 143 - 156, 2010. 6
- [ 21 ] quattroni Piedra 和和。。。  
recognizing 室内的场景。  
在 IEEE 计算机学会计算机会议 2009  
视觉和模式识别 (cvpr 六月 20 - 25,  
2009)
- 2009, 迈阿密, 佛罗里达州, 美国,  
413 - 420 页, 2009. 1、5、6
- [ 22 ]。Reddy 和 M. 沙阿。recognizing  
50 人。  
类别的视频网站。马赫数。可见。  
appl., 24 (5): 971 - 981,  
2013 年。1、5、6
- [ 23 ]。sadanand 和 J. J. 科索。行  
动和高水平的银行表示:  
在视频的活动。在 IEEE 会议的程序  
他是计算机视觉和模式识别, 2012。  
6, 7
- [ 24 ]和 razavian azizpour。谢里夫,  
H, J. 沙利文和 carlsson 页。  
CNN 的特点: 一个是从书架上的  
astounding 基线  
识别。在 IEEE 计算机视觉会议  
模式识别和 workshops 2014 (cvpr),  
六月。6
- [ 25 ]太阳和 J. J. 庞塞。合作学习  
discriminative detectors  
分类和 cosegmentation 的形象。在对  
2013 年的 IEEE 计算机视觉国际会议,  
作者 3400 - 3407, 2013 年。6
- [ 26 ] snchez, J. F. perronnin, T.  
mensink 维贝克, 和 J. J. 的形象  
费舍尔向量的分类: 理论和实践。  
计算机视觉国际杂志, 105 222 - 245,  
(3):  
2013 年。6
- [ 27 ]的土耳其人和 Pentland 和 M。。。  
一系列的特征脸的识别。J。  
cognitive neuroscience, 3 (1): 71  
- 86, 1991. 1
- (29) H. M. A. 王 Ullah 和拉普捷夫  
klaser, 即, C 先生。  
评价时空时空局部特征识别的行动。  
BMVC 在 2009 年英国机器视觉会议  
124.1 - 124.11 页, 2009 年。7
- (30) J. Wang, J. Yang, F. K. 于 LV,  
T. S. Y. 黄, 和工业。  
地方性约束线性编码的图像分类。  
在 3360 - 3367 CVPR, Pages, 2010 年。  
5, 6
- (31) J. Wright, A. Y. Yang, A. P.  
Ganesh Sastry 和 Y. Ma。  
通过稀疏表示的人脸识别的鲁棒性。  
IEEE  
反式。模式分析。马赫。intell., 31  
(2): 210 - 227, 2009 年。1、  
2, 3, 4, 5, 7
- (32) J. Y. K. 宇阳宫, 和 T. S. 黄。  
线性空间  
金字塔匹配使用稀疏编码的图像分  
类。  
在 1794 - 1801 CVPR, 页, 2009 年。5,  
6
- (33) M. L. 杨、张、冯、张 X, D. Fisher  
判别  
词典学习和稀疏表示的。在  
iccv, 页面 543 - 550, 2011 年。2
- (34) M. L. 杨、张、冯、张 X, D. 稀  
疏表示  
基于 Fisher 的歧视字典学习算法  
图像分类。国际计算机视觉学报,  
决策 1 - 24 页, 2014 年。1, 2, 5, 7
- [ 35 ] D. O. C. hamsici, 和你, mart  
mez A. M.。内核优化  
在 discriminant 分析。IEEE 跨。模式  
分析。  
马赫。intell., 33 (3): 631 - 638,  
2011 年。1
- [ 36 ]和 R. M. D. zeiler 弗格斯。  
可视化和理解

卷积网络。在计算机视觉 ECCV 2014  
页面 818 - 833, 2014 年。6  
[ 37 ]。B. 张和李。为 discriminative  
K-SVD 词典  
学习在人脸识别。在 2691 - 2698 CVPR,  
页面,  
2010 年。1, 2, 5, 7  
[ 38 ]周勇沈鹏, J., J. 和风扇。学  
习互相关  
视觉对象识别字典。在 CVPR, 页  
3490 - 3497, 2012 年。2, 5, 7  
(39) 邹 Hastie 和 T. H.。正则化和  
变量选择  
以弹网。英国皇家统计学会杂志》,  
B 系列: 301 - 320 67, 2005 年。2

