

基于密度导向的树型结构核的图像目标分类

陈海林 吴秀清 胡俊华 李斌

(中国科学技术大学电子工程与信息科学系,合肥 230027)

摘要 针对学习算法不能很好地适用于由不等势(unequal cardinality)且无序的特征点集组成的实例,提出密度导向的树型结构核函数。将特征点集自适应分解成一个树型结构,把两个特征点集嵌入该树型结构,形成两个多分辨率直方图。然后计算由这两个多分辨率直方图公共节点所含特征点的密度加权的交叉函数值。该核函数可以自适应确定特征点集之间的局部对应(partial corresponding)关系,具有与特征点数成线性关系的计算复杂度,且是正定的。将该核函数嵌入基于核的判别分类器进行图像目标分类,并与词汇导向的金字塔匹配核进行比较,实验结果表明,密度导向的树型结构核函数能获得较好的分类性能。

关键词 特征点集 树型结构 局部匹配 支持向量机

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2009)12-2545-07

Density-guided Tree-structured Kernel for Image Object Classification

CHEN Hai-lin, WU Xiu-qing, HU Jun-hua, LI Bin

(Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027)

Abstract Density-guided tree-structured kernel is proposed for the situation that the learning algorithm is not fit well for instances consisting of unordered feature sets with unequal cardinality. The feature set is automatically decomposed into a tree, and two feature sets are embedded into this tree to form two multi-resolution histograms, and then the histogram intersection, weighted by the density of feature points in common nodes from two multi-resolution histograms, is computed. The partial correspondences between feature sets can be determined automatically through this kernel, its computation is linear with the number of features, and it is positive-define. This kernel is embedded into kernel based discriminative classifier for image object classification, and compared with vocabulary-guided pyramid match kernel. The experiments show that the density-guided tree-structured kernel can obtain the better classification performance.

Keywords feature sets, tree structure, partial matching, support vector machine

1 引言

在计算机视觉领域,具有各种不变特性的局部特征广泛用于目标匹配或识别^[1-2],图像被描述为无序的特征点集^[3],有些学者研究采用最近邻分类器^[4-5]和基于投票策略的分类器^[3,6]进行目标识别,但这些方法不适用于大尺度训练集。Bosch等人提出基于生成和判别学习联合模型的场景识别,该方

法需要用户指定视觉单词数和主题(topic)数^[7]。基于核函数的学习算法具有优越的学习性能,被广泛运用于模式识别和人工智能领域,如判别分类、回归等等。然而传统的核函数只能适用于固定长度的特征向量,图像必须被描述为一个全局性的特征。例如,Odone等人采用直方图作为整幅图像的特征,用支持向量机(SVM)进行图像分类^[8]。为了较好地学习由局部特征点集组成的图像实例,已提出一些核函数,但这些核函数均具有以下一个或多个缺

收稿日期:2008-07-14;改回日期:2008-10-29

第一作者简介:陈海林(1982~),男。现为中国科学技术大学信号与信息处理博士研究生。主要研究方向为图像理解。

E-mail:chlallex@mail.ustc.edu.cn

点:计算复杂度高、需要建立参数分布(不能很好地描述数据特征)、只适用于含有相等特征点数的特征点集、不满足正定条件、不能很好地建立特征点集所含特征点的局部对应匹配关系、不能适用于高维特征空间、需要人工干预等。

Wolf 和 Shashua 提出了一种具有三次多项式计算复杂度的核函数^[9],而且只对含有相等特征点数的特征点集满足 Mercer 条件;Boughorbel 等人提出了一种非 Mercer 核函数^[10]用于目标识别;Kondor 和 Jebara 首先假设特征点集满足一个参数分布,从而建立一个参数模型^[11],此方法计算量大且适用性不强;Lyu 等人提出了一种具有二次多项式计算复杂度的 Mercer 核^[12],该核函数对大尺度的特征集计算量较大;Grauman 等人提出了金字塔匹配核(PMK)^[13],该核函数首先将特征点集映射到一个多分辨率直方图,然后计算加权的直方图交叉。与其他核函数相比,金字塔匹配核函数具有优越的性能,但只能适用于低维特征空间,最近,Grauman 等人针对此问题又提出了一个词汇导向的金字塔匹配(VGPMK)^[14],但该核函数需要根据经验事先设定金字塔的层数和节点的分支数,没有充分利用特征点集在特征空间中的结构特性。Lazebnik 等人借鉴金字塔匹配核的思想,提出空间金字塔匹配核^[15],该方法要求图像中目标比较规范,需要用户指定空间金字塔层数和通道数。

本文提出密度导向的树型结构核函数(DGTSK)用于学习局部特征集。该方法首先通过局部特征矢量量化在特征空间自适应构造出一棵树,然后将两个特征点集嵌入该树,形成两个多分辨率直方图,再计算这两个直方图的公共节点所含特征点的密度加权的直方图交叉,从而得到密度导向的树型结构核函数。通过在特征空间自适应建立树型结构和估计公共节点包含的特征点密度,可以充分地挖掘特征点集在局部特征空间的结构特性,能够很好地提升分类性能。

与已有核函数相比,DGTSK 同时具有以下优点,是非参数核函数,具有与特征点数目呈线性关系的计算复杂度,能够计算出具有不等势(unequal cardinality)的两个特征点集之间局部匹配(partial matching)关系,在高维特征空间也具有有良好的匹配性能,无需用户指定特定参数,且是正定的、满足 Mercer 条件,可以用于基于核函数的学习算法,能够将图像目标的描述和分类器进行良好的融合,进行

图像目标定位或识别。本文运用 DGTSK 在公共数据库 ETH-80^[13]上进行局部匹配实验,并与最优局部匹配算法的一种实现方法^[16]进行比较,结果表明 DGTSK 能够较好地近似最优局部匹配算法,具有良好的局部匹配性能。还将 DGTSK 嵌入到 SVM,在 ETH-80 上进行目标分类实验,实验结果表明其分类精度优于 VGPMK,且具有较好的稳定性。

2 原理和方法

图像分类、目标识别问题需要对两幅图像的局部区域建立匹配关系,以便量化目标之间的相似度或在图像中定位目标^[5,13],这是符合人的视觉观点,在理解图像内容时人需要在图像之间建立目标或语义内容的匹配关系。采取有效方式确定特征点集之间的局部匹配关系对局部特征点集的学习至关重要。

下面介绍特征点集之间理论上最优的局部匹配基本原理^[14],然后介绍本文的局部对应匹配原理以及核函数的形成方法。

设 d 维特征空间 $F, F \subseteq \mathbb{R}^d$, 输入空间 I 包含了 F 中的一系列特征点集, $I = \{X \mid X = \{x_1, \dots, x_s, \dots, x_r\}\}, x_s \in F$, 其中 $r = |X|$ 表示 X 的势,即 X 中的特征点数,且 r 可以随 I 中的元素不同而不同。

设两个特征点集 $Y \in I$ 和 $Z \in I, g = |Y|, h = |Z|$, 且 $g \leq h$, 则 Y 和 Z 之间的局部匹配 $\varphi(Y, Z; \pi) = \{(y_1, z_{\pi_1}), \dots, (y_g, z_{\pi_g})\}$ 是将 Y 中每个点匹配到 Z 中的一个点,与 Y 中特征点相对应的 Z 中特征点索引为

$$\pi = [\pi_1, \dots, \pi_i, \dots, \pi_g], 1 \leq \pi_i \leq h \quad (1)$$

式中, π_i 指定 $z_{\pi_i} \in Z$ 对应 $y_i \in Y, 1 \leq i \leq g$, 局部匹配的代价函数为

$$C(\varphi(Y, Z; \pi)) = \sum_{y_i \in Y} \|y_i - z_{\pi_i}\| \quad (2)$$

$$\pi^* = \arg \min_{\pi} C(\varphi(Y, Z; \pi)) \quad (3)$$

理论上最优局部匹配 $\varphi(Y, Z; \pi^*)$ 满足式(3)。文献[16]针对最优局部匹配算法提出了一种基于线性规划的实现方法,至少具有多项式时间的计算复杂度,在最坏情况下具有指数时间的计算复杂度,很难在实际中应用。最优匹配算法只能建立一对一的对应关系,不能建立一对多或多对多的对应关系,不能有效地提取局部对应域,不能很好地用于图像

目标的匹配和分类等。

本文提出密度导向的树型结构核,可以隐式确定特征点集之间的局部匹配关系,很好地学习局部特征点集。首先利用一系列图像的局部特征点集在特征空间自适应构建树型结构,然后将两幅图像的局部特征嵌入该结构,形成两个多分辨率直方图,计算这两个多分辨率直方图的公共节点中特征点匹配数,采用所包含特征点密度加权相应的特征点匹配数,从而得到一个局部匹配代价。利用两个多分辨率直方图的公共节点所含特征点的匹配,可以对两个特征点集从树根到树叶建立由粗到细的局部对应关系。

2.1 自适应构建树型结构

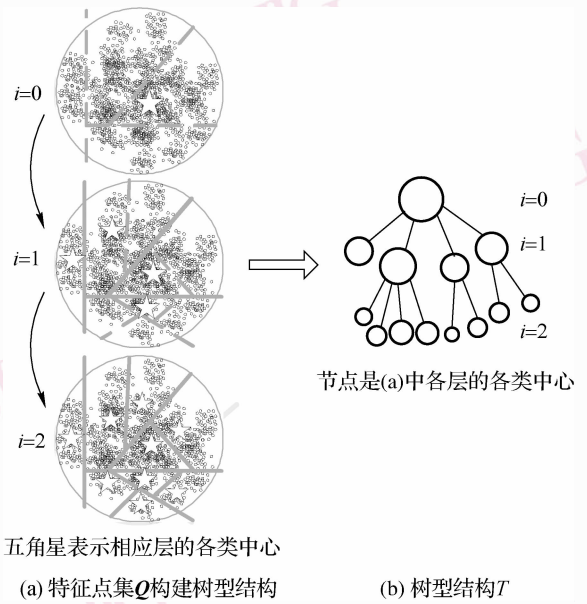
选择一系列特征点集 Q 来构建树型结构,首先将 Q 的中心作为树的根 ($i=0$),根据 Q 中特征点的方差分析其在特征空间的分布情况,由所有维的方差组成一个方差向量 $\nu = (\nu_1, \nu_2, \dots, \nu_d)$,按式(4)求得归一化向量 ν^* , ν^* 中元素 $\nu_k^*, k=1, \dots, d$,大于 $1/\sqrt{d}$ 的个数为 c_{00} ,则将 Q 用 K 均值聚成 $2^{c_{00}}$ 类,作为树根的分支(儿子)节点,得到 $2^{c_{00}}$ 个聚类中心。然后按照上述方法对每一分支节点进行递归聚类,当到达一定层数且节点的自适应分支数大于某阈值(可以根据分支节点中所含特征点数确定阈值),将该节点作为树的一个叶节点,当所有节点都递归完,则形成一棵树 T ,它的节点是各聚类中心点,如图 1 所示,递归聚类的过程可以为树的最大层数设置一个阈值,以防止某些“畸形”特征点影响树的结构,所谓“畸形”特征点是选择的特征点集中出现的不常见点。

$$\nu^* = (\nu_1^*, \nu_2^*, \dots, \nu_d^*) \quad (4)$$

式中, $\nu_k^* = \frac{\nu_k}{\|\nu\|}, k=1, \dots, d$ 。

由树型结构 T 的构造方法可知, T 的第 i 层第 j 个节点的儿子数为 $2^{c_{ij}}, 0 < c_{ij} < d, 0 \leq i < L-1, 0 \leq j \leq M_i-1, L$ 为树 T 的最大层数, M_i 为 T 的第 i 层的总节点数。当 d 较大时可能导致 $2^{c_{ij}}$ 很大,可根据 d 值将 $2^{c_{ij}}$ 线性映射到一定范围内。

构建树型结构 T 后,将 I 的特征点集 $X \in I$ 嵌入到 T 中,形成多分辨率直方图 $T(X)$,方法如下:将 X 中每一个点按照父子关系沿着 T 的根到叶节点,计算该点到 T 的每层相应节点的距离,距离最小的节点就是该点在本层的位置。也就是说一个点首先被分配给树根,然后被分配给树根的一个儿子节点,



五角星表示相应层的各类中心

(a) 特征点集 Q 构建树型结构

(b) 树型结构 T

图 1 构建树型结构的过程

Fig. 1 Process for building the tree structure

再被分配给该子节点的其中一个儿子,一直递归下去,直到叶节点。将 X 嵌入到 T 中形成 $T(X)$, $T(X) = [T_0(X), \dots, T_i(X), \dots, T_{L-1}(X)], 0 \leq i \leq L-1, L$ 为树的最大层数。其中, $T_i(X) = [\langle P, n \rangle_1, \dots, \langle P, n \rangle_j, \dots, \langle P, n \rangle_{N_i}]$, 表示 N_i 维直方图, $1 \leq j \leq N_i$, 当 $i > 0$ 时, N_i 表示 T 的第 $i-1$ 层的儿子数, $N_0 = 1$, $\langle P, n \rangle_j$ 中的 P 表示特征点按照父子关系从树根分配到第 i 层第 j 个节点的路径,称为路径向量,可表示为 $[p_0, \dots, p_q, \dots, p_i], q < i, p_q \in [0, r_q - 1], r_0 = 1$, 当 $q > 0$ 时, r_q 表示第 $q-1$ 层各节点最大儿子数。 $\langle P, n \rangle_j$ 中的 n 表示将特征点集 X 嵌入到树型结构 T 后, X 中的点被嵌入到 T 的第 i 层第 j 个节点的点数。

2.2 多分辨率直方图公共节点的密度估计方法

由下面的公共节点密度估计原理可以看出,利用两幅图像的局部特征点集所构成的两个多分辨率直方图的公共节点所含特征点的密度来加权直方图交叉,可以不受大特征集中“出格点”的干扰,能充分挖掘两幅图像所含特征点集在特征空间上结构的相似性。

2.2.1 理论上最优密度估计原理

两特征点集 $Y \in I$ 和 $Z \in I$ 嵌入到 T , 形成两个多分辨率直方图 $T(Y)$ 和 $T(Z)$ 。用 O_{ij} 表示 $T(Y)$ 和 $T(Z)$ 的公共节点,即式(10)中 $n_{ij}(Y)$ 和 $n_{ij}(Z)$ 均不为零,是 T 的第 i 层第 j 个节点,其分布示意图

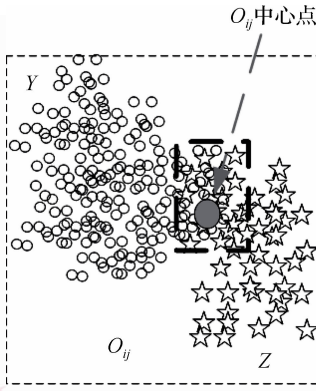


图 2 特征点集 Y 和 Z 嵌入 T 后公共节点 O_{ij} 的分布示意图

Fig. 2 Illustration for distribution of the common node O_{ij} , when feature sets Y and Z are embedded into T

如图 2 所示,里面的虚线框表示 Y 和 Z 在 O_{ij} 中特征点在特征空间中的重叠区域,记为 Z_{ij} ,则节点 O_{ij} 的密度 ρ_{ij} 为 Z_{ij} 所含特征点的密度。 ρ_{ij} 的估计可以采取参数密度估计方法,如 Mean Shift^[17],但该方法计算量较大,较复杂。

2.2.2 近似密度估计方法

在实际中准确计算出重叠区域 Z_{ij} 计算量较大,尤其是在高维特征空间。所以在实际运用时采用简单有效的近似密度估计方法,具体的估计可以采取如下 3 种方法。

(1) 计算两特征点集 Y 和 Z 在 O_{ij} 中的点到 O_{ij} 中心点的平均距离为 $d_{m_{ij}}$,则 O_{ij} 的密度为

$$\rho_{ij} = e^{-d_{m_{ij}}} \quad (5)$$

$$\text{或 } \rho_{ij} = \frac{1.0}{1.0 + d_{m_{ij}}} \quad (6)$$

$T(Y)$ 和 $T(Z)$ 的非公共节点的密度为 0。

(2) 计算 $T(Y)$ 和 $T(Z)$ 在公共节点 O_{ij} 的子特征点集的中心点之间距离 $d_{b_{ij}}$,则 O_{ij} 的密度为

$$\rho_{ij} = e^{-d_{b_{ij}}} \quad (7)$$

$$\text{或 } \rho_{ij} = \frac{1.0}{1.0 + d_{b_{ij}}} \quad (8)$$

$T(Y)$ 和 $T(Z)$ 的非公共节点的密度为 0。

(3) 设 $T(Y)$ 和 $T(Z)$ 在公共节点 O_{ij} 中总点数为 A , x_r 为 $T(Y)$ 和 $T(Z)$ 在 O_{ij} 中的点, $x_{o_{ij}}$ 表示 O_{ij} 的中心点,则

$$\rho_{ij} = \frac{1}{A} \sum_{r=0}^{A-1} e^{-\|x_r - x_{o_{ij}}\|} \quad (9)$$

2.3 密度导向的树型结构核函数

设 $n_{ij}(Y)$ 和 $n_{ij}(Z)$ 分别为 $T_i(Y)$ 和 $T_i(Z)$ 中 $\langle P, n \rangle_j$ 的 n , 密度导向的树型结构核函数为

$$\zeta(T(Y), T(Z)) = \sum_{i=0}^{L-1} \sum_{j=0}^{M_i-1} \rho_{ij} \min(n_{ij}(Y), n_{ij}(Z)) \quad (10)$$

式中, $\min(n_{ij}(Y), n_{ij}(Z))$ 表示特征点集 Y 和 Z 嵌入 T 后在第 i 层第 j 个节点的特征点匹配数,表示 Y 和 Z 中点的局部匹配关系。如果 $n_{ij}(Y)$ 和 $n_{ij}(Z)$ 均为 0, 则 ρ_{ij} 为 0。综合考虑计算复杂度和核函数性能, ρ_{ij} 按式(6)计算,该方法计算简单,因为特征点集嵌入 T 时需要计算特征点集中所含点到 O_{ij} 中心点的距离。输入特征点通常是归一化向量,每个层次节点均含有一定数量的特征点,根据式(6)可知,不同层次节点的 ρ_{ij} 不会出现较大的量级上的差异。

2.4 证明 DGTSK 满足 Mercer 条件

基于核函数的学习算法通过非线性映射 Φ 将低维特征空间 E_L 中数据映射到高维特征空间 E_H , 即 $\Phi: E_L \rightarrow E_H$, 然后在高维特征空间寻找线性最优分类平面。只有正的半定 (positive semi-definite) 核函数才能保证基于核函数的凸优化学习算法 (如支持向量机 SVM) 寻找到最优分类平面。

直方图交叉函数是满足 Mercer 条件、正定的^[3], 所以直方图 $T_i(Y)$ 和 $T_i(Z)$ 的各节点的交叉 $\min(n_{ij}(Y), n_{ij}(Z))$ 是 Mercer 核。由于 Mercer 核函数对非负线性联合封闭, 根据式(10)可知其满足 Mercer 条件。所以密度导向的树型结构核函数能被运用于任何需要 Mercer 核的学习算法。

2.5 局部匹配性能分析

计算两个特征点集密度导向的树型结构匹配核函数的过程,反映了各特征点集内部点之间的依赖关系,同时也能够获得两个特征点集之间的对应关系。通过对多分辨率直方图公共节点中所含特征点的密度估计,能良好地挖掘两特征点集在特征空间上局部匹配程度。DGTSK 可以用于两个不等势的特征点集,将较小势的特征点集匹配到具有较大势的特征点集的子集,仅计算特征点集最相似的子集之间相似度,而忽略其他非相似的特征点的影响,这可以减少图像背景、目标位置、目标亮度和目标形状等变化对目标分类性能的影响, DGTSK 具有较强的局部对应匹配能力。

2.6 计算复杂度分析

构建特征点集 X 的多分辨率直方图 $T(X)$ 的所需计算量为 $O\left(dm_X \sum_{i=0}^{L-1} H_i\right)$, $m_X = |X|$, $H_0 = 1$, 当 $L > i > 0$ 时, H_i 表示 T 的第 $i-1$ 层所有节点的儿子

数的最大值, d 为特征维数。利用 $T(\mathbf{Y})$ 和 $T(\mathbf{Z})$ 计算核函数所需计算量为 $O\left(\sum_{i=0}^{L-1} \min(m_y, m_z, M_i)\right)$, 其中, $m_y = |\mathbf{Y}|, m_z = |\mathbf{Z}|, M_i$ 为 T 的第 i 层总节点数。可见, 多分辨率直方图的构建以及核函数的计算所需计算复杂度均与特征点集所含特征点数和 d 呈线性关系。

下面分析 VGPMK 的计算复杂度, 设 VGPMK 的金字塔层数和分支数分别为 L_v 和 B 。特征点集 \mathbf{X} 嵌入到金字塔所需计算量为 $O(dm_x BL_v)$ 。计算词汇导向的金字塔匹配核函数所需计算量为 $O(L_v \min(m_y, m_z))$ 。

可见, VGPMK 与 DGTSK 的计算复杂度一样, 都具有线性计算复杂度。

3 实验与结果分析

实验数据为 ETH-80^[13], 该数据库共含有 8 类目标, 分别为苹果, 车, 牛, 杯子, 狗, 马, 梨, 西红柿, 如图 3 所示。每一类有 10 个目标, 每个目标从不同位置、不同视角拍摄 5 幅图像, 实验总共 400 幅图像。均匀提取 SIFT (scale invariant feature transform) 特征^[3], 每幅图像含有约 256 个 128 维 SIFT 特征点。

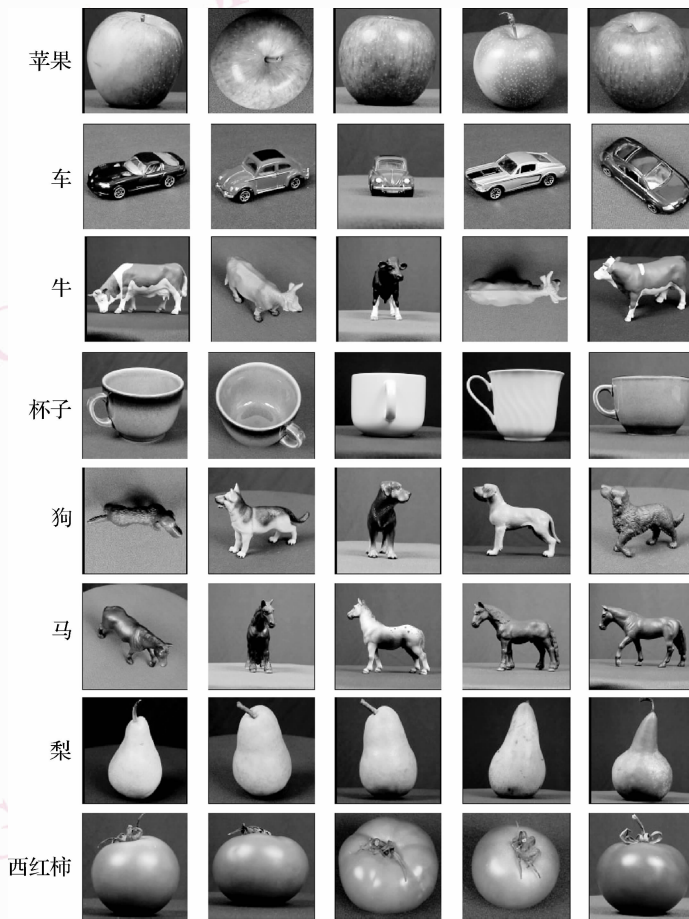


图 3 ETH-80 数据库图像举例

Fig. 3 Example images from ETH-80 database

3.1 局部匹配实验与性能分析

对提取的 128 维 SIFT 用主成分分析 (PCA) 获取不同维数的局部特征。从图像库中选取 300 幅图像用于构建树型结构, 其余 100 幅图像用于测试部分匹配性能, 采用 Spearman 排序相关系数 $R^{[14]}$ 度量 DGTSK 和 VGPMK 对最优匹配算法^[16] 的近似程度,

$$R = 1 - 6 \sum_{i=1}^N D^2 / N(N^2 - 1) \quad (11)$$

式中, D 表示排序之间的差别, N 表示用于排序的元素个数。DGTSK 和 VGPMK 在不同维数的特征空间与最优匹配算法的 Spearman 相关系数如图 4 所示, 图中表示 5 次实验的平均值, 其中 VGPMK 将两个特

征点集中的点到金字塔节点中心的最大距离作为金字塔直方图交叉的加权系数。

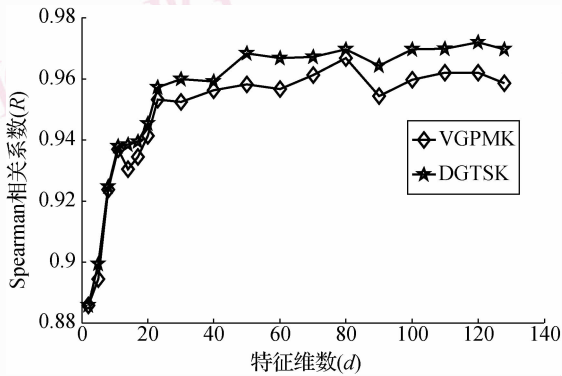


图 4 DGTSK 和 VGPMK 与最优匹配算法的 Spearman 相关系数

Fig. 4 Spearman correlation coefficients for DGTSK and VGPMK with optimal match

从图 4 中可以看出,随着特征空间维数的增加, DGTSK 和 VGPMK 可以保持良好的匹配精度, DGTSK 的部分匹配性能要略优于 VGPMK,且无需用户指定参数。

3.2 目标分类实验与结果分析

将 DGTSK 嵌入 SVM 在 ETH-80 上进行目标分类实验,采用交叉验证 (cross-validation) 的方式评价分类性能。对每幅图像提取的 128 维 SIFT 特征,用 PCA 得到 10 维局部特征,运用 VGPMK 和 DGTSK 对 128 维和 10 维局部特征进行目标分类实验。每一类中选取一个不同目标的 5 幅图像作为测试图像,其余作为训练图像,结果如表 1 所示,表中分类结果是所有目标的平均分类精度。每次学习和测试所用的图像在数据集中的比例分别为 395 : 400 和 5 : 400。实验时 VGPMK 的金字塔层数和每一节点的分支数分别取 5 和 10。由表 1 可以看出, DGTSK 的分类性能要明显优于 VGPMK。在高维特征空间, DGTSK 对 ETH-80 的分类也取得了较好效果,但相对于低维空间,分类精度有所降低,这可能是因为在高维空间目标背景对分类产生了较大的混淆。

表 1 ETH-80 目标分类结果

Tab. 1 ETH-80 object classification result

单位: %

核函数	平均分类精度	
	$d = 10$	$d = 128$
VGPMK	78.25	62.50
DGTSK	84.00	80.75

DGTSK 在 10 维特征空间上对 ETH-80 的具体分类结果如图 5 所示,其中,第 i 行第 j 列表示第 i 类误分为第 j ($j \neq i$) 类的比率,等于 0 的数字未在表中标出,行元素之和为 1,对角线上元素代表相应类的正确分类率,对角线上元素的平均值表示平均正确分类率,为 84.00%。可以看出,运用 DGTSK 在分类时对苹果和西红柿较容易产生混淆,对狗、牛和马也较容易产生混淆,车和梨与其他类的目标在视觉上有较大差异,获得了百分百的正确率,这些符合人的视觉观点。

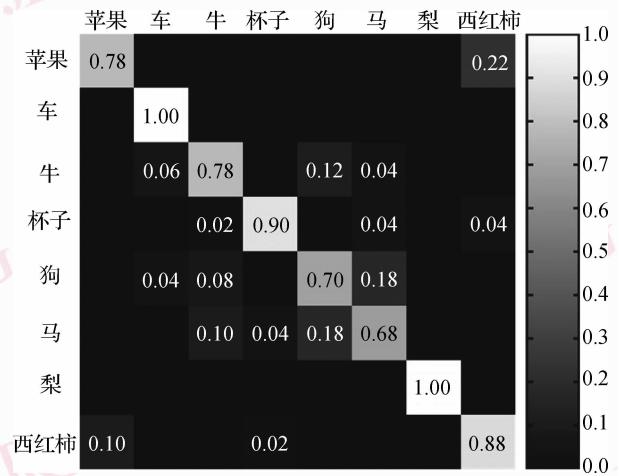


图 5 DGTSK 的分类结果混淆表

Fig. 5 Classification confusion table for using DGTSK

采用基于 VGPMK 的两个函数在 10 维特征空间对 ETH-80 进行分类实验,方法 1 (VGPMK-Input-Specific) 采用文献 [14] 中的加权方式,方法 2 (VGPMK-Density) 采用多分辨率直方图的公共节点所含特征点的密度对 VGPMK 进行加权。平均分类结果如图 6 所示,图 6 中 (a) 表示固定金字塔分支数为 5,改变金字塔层数, (b) 表示固定金字塔层数为 3,改变金字塔分支数,可以看出,只有在金字塔层数和分支数均为 5 的情况下, VGPMK-Density 与 VGPMK-Input-Specific 表现出了相同的分类性能,而在其他情况下, VGPMK-Density 的平均分类率要比 VGPMK-Input-Specific 高出几个百分点,可见,多分辨率直方图的公共节点的密度加权对分类性能具有良好的提升作用。

4 结论

针对由不等势且无序的特征点集组成的图像目标的学习问题,提出了密度导向的树型结构核函数,

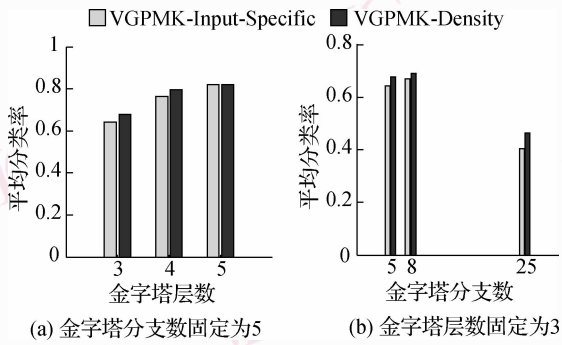


图 6 公共节点的密度加权对分类性能的影响

Fig. 6 Effect to classification performance from weighting by density of feature points in the common nodes

分析局部特征集在特征空间的分布特征,将特征空间自适应地表示为一个树型结构,通过估计两个多分辨率直方图的公共节点所含特征点的密度来分析局部特征点集在特征空间的相似性,用此密度加权直方图交叉,形成密度导向的树型结构核。该核函数能充分挖掘特征点集在特征空间中的结构特性,将其嵌入 SVM 对 ETH-80 进行分类实验,结果表明该核函数具有良好的分类能力。通过局部匹配实验可以看出,密度导向的树型结构核函数能够对理论最优的局部匹配算法进行较好地近似,具有良好的局部对应能力,可以应用于目标匹配和目标分类等。

参考文献 (References)

- 1 Wang Jun-qiu, Zha Hong-bin. Combining interest points and edges for building and object recognition [J]. *Journal of Computer-aided Design & Computer Graphics*, 2006, **18**(8): 1257-1263. [王君秋, 查红彬. 结合兴趣点和边缘的建筑物和物体识别方法[J]. *计算机辅助设计与图形学报*, 2006, **18**(8): 1257-1263.]
- 2 LIU Bo-wen, YU Song-yu, XU Yi, et al. Correspondence of object-of-interest in wide baseline active vision [J]. *Journal of Image and Graphics*, 2007, **12**(10): 1917-1921. [刘博文, 余松煜, 徐奕等. 宽基线主动视觉中感兴趣目标的对应技术[J]. *中国图象图形学报*, 2007, **12**(10): 1917-1921.]
- 3 Lowe D. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, **60**(2): 91-110.
- 4 Belongie S, Malik J, Puzicha J. Shape matching and object

- recognition using shape contexts [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(24): 509-522.
- 5 Berg A, Berg T, Malik J. Shape matching and object recognition using low distortion correspondences [A]. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]*, San Diego, USA, 2005: 26-33.
- 6 Mikolajczyk K, Schmid C. Indexing based on scale invariant interest points [A]. In: *Proceedings of IEEE Conference on Computer Vision [C]*. Vancouver, Canada, 2001: 525-531.
- 7 Bosch Anna, Zisserman Andrew, Munoz Xavier. Scene classification using a hybrid generative/discriminative approach [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(4): 712-727.
- 8 Odone F, Barla A, Verri A. Building kernels from binary strings for image matching [J]. *IEEE Transactions on Image Processing*, 2005, **14**(2): 169-180.
- 9 Wolf L, Shashua A. Learning over sets using kernel principal angles [J]. *Journal of Machine Learning Research*, 2003, **4**(1): 913-931.
- 10 Boughorbel S, Tarel J-P, Fleuret F. Non-mercer kernels for SVM object recognition [A]. In: *Proceedings of British Machine Vision Conference [C]*, London, UK, 2004: 137-146.
- 11 Kondor R, Jebara T. A kernel between sets of vectors [A]. In: *Proceedings of International Conference on Machine Learning [C]*, Washington, D C, USA, 2003: 361-368.
- 12 Lyu S. Mercer kernels for object recognition with local features [A]. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]*, San Diego, USA, 2005: 223-229.
- 13 Grauman K, Darrell T. The Pyramid match kernel: discriminative classification with sets of image features [A]. In: *Proceedings of IEEE Conference on Computer Vision [C]*, Beijing, China, 2005: 1458-1465.
- 14 Grauman K, Darrell T. Approximate correspondences in high dimensions [A]. In: *Proceedings of Advances in Neural Information Processing Systems (19) [C]*, Cambridge, Massachusetts, USA: MIT Press, 2007: 505-512.
- 15 Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories [A]. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]*, New York, USA, 2006: 2169-2178.
- 16 Rubner Y, Tomasi C, Guibas L. The earth mover's distance as a metric for image retrieval [J]. *International Journal of Computer Vision*, 2000, **40**(2): 99-121.
- 17 Dorin Comaniciu, Peter Meer. Mean Shift: A robust approach toward feature space analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603-619.