

多尺度局部结构主导二值模式学习图像表示

张东波* 易良玲 许海霞 张莹

(湘潭大学信息工程学院 湘潭 411105)

(机器人视觉感知与控制国家工程实验室 长沙 410012)

摘要: 通过零均值化的微观结构模式二值化(ZMPB)处理, 该文提出一种立足于局部图像多尺度结构二值模式提取的图像表示方法。该方法能够表达图像中可能出现的各种具有视觉意义的重要模式结构, 同时通过主导二值模式学习模型, 可以获得适应于图像数据集的主导特征模式子集, 在特征鲁棒性、鉴别力和表达能力上达到优异性能, 同时可以有效降低特征编码的维度, 提高算法的执行速度。实验结果表明该算法性能优异, 具有很强的鉴别能力和鲁棒性, 优于传统LBP和GIMMRP方法, 和很多最新算法结果相比, 也具有竞争优势。

关键词: 目标识别; 零均值化的微观结构模式二值化; 主导二值模式学习; 局部结构

中图分类号: TP391.4

文献标识码: A

文章编号: 1009-5896(2019)04-0896-08

DOI: 10.11999/JEIT180512

Multi-scale Local Region Structure Dominant Binary Pattern Learning for Image Representation

ZHANG Dongbo YI Liangling XU Haixia ZHANG Ying

(College of Information Engineering, Xiangtan University, Xiangtan 411105, China)

(Robot Visual Perception & Control Technology National Engineering Laboratory, Changsha 410012, China)

Abstract: By means of Zero-mean Microstructure Pattern Binarization (ZMPB), an image representation method based on image local microstructure binary pattern extraction is proposed. The method can express all the important patterns with visual meaning that may occur in the image. Moreover, through the dominant binary pattern learning model, the dominant feature pattern set adapted to the different data sets is obtained, which not only achieves excellent ability in feature robustness, discriminative and representation, but also can greatly reduce the dimension of feature coding and improve the execution speed of the algorithm. The experimental results show that the proposed method has strong discriminative power and outperforms the traditional LBP and GIMMRP methods. Compared with many recent algorithms, the proposed method also presents a competitive advantage.

Key words: Object recognition; Zero-mean Microstructure Pattern Binarization (ZMPB); Dominant binary pattern learning; Local region structure

1 引言

计算机对图像进行理解通常采用“从局部到全局”的策略, 这是一个从微观到介观到宏观由底向上的图像描述过程, 由于单一像素点本身不能表示

任何有意义的视觉结构, 能够观测到有意义视觉结构的最小区域是 3×3 的窗口邻域, 其中相邻像素点之间的灰度差异会表现为特定的低层视觉结构, 所以本文所说的微观尺度指的就是图像中 3×3 的窗口邻域; 介观尺度表示图像的局部区域, 代表图像中可能出现的部件; 宏观尺度则表示整个图像或物体。

传统的局部特征提取技术通常基于局部区域描述, 例如SIFT^[1]、SURF^[2]描述关键点邻域的 4×4 格状空间, GLOH^[3]将邻域划分为16扇区, DAISY^[4]采用圆形邻域, HOG^[5]利用细胞单元区域等。除此以外, 大家熟知的最早的基于微观邻域结

收稿日期: 2018-05-25; 改回日期: 2018-12-18; 网络出版: 2018-12-25

*通信作者: 张东波 zhadonbo@163.com

基金项目: 国家自然科学基金(61602397), 湖南省自然科学基金(2017JJ2251, 2017JJ3315), 湖南省重点学科建设项目

Foundation Items: The National Natural Science Foundation of China (61602397), The Natural Science Foundation of Hunan Province (2017JJ2251, 2017JJ3315), The Key Discipline Construction Project of Hunan Province

构描述的是局部二值模式(LBP)算子。LBP通过比较中心像素和邻域像素的灰度值大小实现微观结构的二值模式编码, LBP^[6]最初定义于像素的8邻域, 为了提取大尺度结构纹理特征, Ojala等人^[7]对LBP进行了改进, 可以在定义的半径为 $R(R > 0)$ 的圆周均匀采样 $P(P > 0)$ 个邻域像素。但是如果采样点过多, 则二值模式数目急剧增加, 为了减少模式数, Ojala等人^[7]提出了LBP统一模式, 即在二进制数进行一次循环运算时, 最多只产生两位变化的模式。为了获得旋转不变性, Ojala等人^[7]又提出了旋转不变LBP, 旋转不变LBP与统一模式LBP联合, 则可以获得旋转不变统一LBP。为了保留图像中曲率较大、交叉点和角点处的模式, Liao等人^[8]提出了显性局部二值模式(Dominant LBP, DLBP), 该方法使用二值模式中出现频率较高的模式来描述纹理信息, DLBP方法对于图像旋转以及图像噪声有良好的适应性。由于LBP方法具有计算复杂度低, 速度快, 同时又具有旋转不变性和灰度不变性等显著优点, 因而该方法被广泛应用于图像匹配, 目标检测与识别等领域。然而, 常规的LBP方法存在以下缺陷: (1) LBP在其二值化过程中只保留了像素间的灰度大小关系, 图像中的信息损失较大; (2) LBP可能产生不均匀分布的直方图和低频模式类型; (3) LBP对所关注的模式类型需要预先定义, 如“统一模式”是一种预先定义的模式。

最近, 文献^[9]提出GIMMRP, 该算子通过直接对图像 3×3 邻域结构进行模板匹配来实现二值编码, 为了找到微观结构中比较稳定的具有明确视觉含义的模式结构, 例如边缘、曲线/直线、角点、曲线/直线端点, 斑点(极值点), 文献^[9]将重要执行二值模式分为类似于LBP的统一模式、线条类、交叉点/分叉点3类总共132种, 而其余模式均被归为混合模式类。GIMMRP不仅具备LBP类似的优点, 而且具备很强的微观结构描述能力, 大大提升了二值描述算子的鉴别能力, 在目标检测和识别领域具有广泛的适用性。但是该算子的特征模式编码过程需要和512种模板进行匹配计算(卷积), 因此特征计算比较耗时, 此外, 考虑到不同的图像数据集中重要执行模式的分布会呈现较大差异, 选取全部512种模式或人为经验筛选的模式子集很难适应不同图像数据本身的特点, 因此, 本文设计算

子避免模板匹配计算, 同时引入机器学习算法, 通过同时考虑特征模式的鲁棒性, 判别能力和表示能力, 对于不同的数据集, 本文能够学习得到不同的稳定的主导模式子集, 并通过空间池化操作实现对图像的整体描述, 在降低特征维数的同时提高算法鲁棒性。

2 方法原理

2.1 基于零均值化的微观结构模式二值化(ZMPB)

由于LBP模式只对邻域周边采样点进行二值编码, 完全排除了中心点, 这一定程度破坏了中心点和邻域点的关联关系, 造成模式类数量的减少, 从而使得部分微观结构不能得到有效表示。 3×3 邻域是图像中能够表示有意义微观结构的最小区域, 其二值化后的模式对应的视觉结构可能是 $2^9 = 512$ 种中的任意一种, 图1给出了部分模式示例, 其中有的代表边缘、角点、端点、极值点、线条、交叉点/分叉点等各种低层视觉结构, 也有部分模式缺乏视觉含义, 属于图像噪声造成的杂乱模式类。

为了对原始灰度图像的 3×3 邻域进行二值模式编码, 文献^[9]基于最大响应原则, 提出一种二值模式匹配编码方法。该方法通过对图像中每一个 3×3 邻域和512种模板进行卷积计算, 将获得最大响应的二值模式作为该像素点邻域的编码。文献^[9]提出的GIMMRP在二值编码过程时保留了中心像素点, 充分利用了 3×3 邻域中全部像素点的关联信息。但是由于每一个 3×3 邻域在提取编码值时, 需要和512个模板进行卷积运算, 因此比较耗时。为了加快计算效率, 本文提出零均值化的微观结构模式二值化(ZMPB)方法, 该方法避免了模板匹配卷积计算, 而且编码结果完全等价于GIMMRP模式。

对图像进行 1×1 分块的ZMPB方法计算示例如图2所示, 对于灰度图像 I 的某 3×3 邻域, 假设中心像素记为 b_0 , 其8邻域的像素点记为 $b_i(i = 1, 2, \dots, 8)$, 如图2(a), $P = 9$ 为区域对应的像素个数, 以中心点作为最高位, 从左上角开始, 沿顺时针方向, 按2的整数次幂分配权重, 如图2(d)所示。

计算 3×3 邻域点灰度平均值 Ave_P :

$$Ave_P = \sum_{i=1}^P b_i / P \quad (1)$$

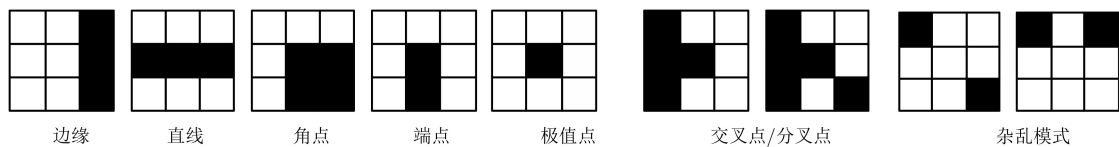


图1 模式示例图

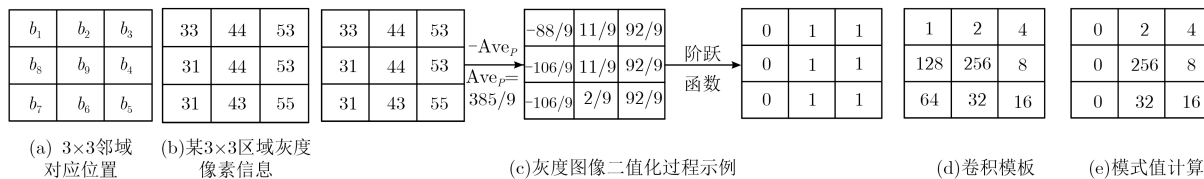


图 2 ZMPB模式计算示例

每个相邻像素点 b_i 减去均值 Ave_p 后，通过符号函数得到 3×3 区域的二值编码信息

$$b'_i = \text{sgn}(b_i - Ave_p) \tag{2}$$

某像素点的 ZMPB 模式值计算示例如图 2(c) — 图 2(e) 所示。101111100 是求得的二值编码模式，其 ZMPB 模式值 = 256 + 32 + 16 + 8 + 4 + 2 = 318。

不难证明，ZMPB 模式和最大响应模式 GIMMRP 是完全等价的。假设给定 3×3 邻域的均值为 m ，和该邻域做卷积运算的 512 个模板构成的集合记作 Ω_s ，则模板集 Ω_s 和 3×3 邻域卷积后响应最大的模式可以表示为

$$\text{CONV_}P_{\max} = \arg \max_{\{a_i | a_i \in \{0,1\}, i=1, \dots, 9\} \in \Omega_s} \sum_{i=1}^9 a_i (b_i - m) \tag{3}$$

由于 $a_i \in \{0, 1\}$ ，显然使式 (3) 取得最大值的匹配模式应该满足：如果 $b_i - m < 0$ ，则 $a_i = 0$ ，否则 $b_i - m > 0$ ，则 $a_i = 1$ 。

按照上述进行判断所获得的具有最大响应的模式正好是对 3×3 邻域进行零均值化处理后进行二值化的结果，所以 ZMPB 模式和最大响应模式 GIMMRP 等价。但是 GIMMRP 编码时每一个像素点编码都需要做 512 次卷积运算 (其中共包含 $512 \times 3 \times 3$ 次乘法计算，512 次累加和计算，512 维的数组的比较找最大值计算)，显而易见编码效率很低。而 ZMPB 编码只要做 1 次零均值化 (含 1 次累加和计算， 3×3 次减法计算) 和 1 次二值化 (3×3 次正负值判断) 操作即可，因此其模式提取时间大大减少，编码效率非常快。

2.2 主导模式学习算法

ZMPB 可能的二值模式总数为 $2^9 = 512$ 种，但是其中只有少部分模式具有明确的视觉含义，通常具有明确视觉含义的微观结构比较稳定，出现频率高，比较重要。而没有明确视觉含义的其它模式通常在大幅噪声的情况下或纹理结构杂乱的背景中才会出现，在常规图像中出现的频率很低。找到其中重要的执行模式，不仅在图像发生变换或干扰时可以实现更为稳定的表达，同时可以获得更好的识别性能。此外，对于不同类型的图像，例如对手写数字图像，车标图像、人脸图像、自然场景图像等进

行描述时，显然有的图像边缘结构信息丰富，有的具有较多的纹理细节或较大的视角和光照变化，因此进行描述时不同类型的图像需要不同的模式子集。鉴于上述情况，本文引入主导模式学习模型，通过学习算法找到适应于不同数据集的最佳模式子集，从而大大降低特征维度，加快算法执行效率。

2.2.1 主导模式学习模型

假定 $X = [x_1, x_2, \dots, x_j] \in R^{d \times N}$ 为训练样本，其中 d 和 N 分别代表输入空间的维数和样本数，假定训练样本包含 C 类，且第 i 类包含 N_i 个样本 ($\sum_{i=1}^C N_i = N$)。为了学到一个期望的性能优良的主导模式集，学习模型中需要考虑多种因素，其中包含特征的鲁棒性，特征的鉴别能力和特征的表示能力 3 个方面，以手写数字为例，其主导模式学习模型如图 3 所示。

特征鲁棒性： 图像中出现频率高的模式比较可靠，反映了图像中比较稳定的视觉结构，因此模式出现频率是衡量特征鲁棒性的重要指标。通过式 (4)，可以找到每个训练图像中出现频率最高的模式组成的主要模式子集，其中 A 是属于类 i 的训练图像 x_j 中每个 ZMPB 模式值出现的次数按降序排列的结果

$$J_j = \arg \min \left(\frac{\sum_{l=1}^h A(j, l)}{\sum_{k=1}^q A(j, k)} \right) \geq n \tag{4}$$

其中， $\sum_{l=1}^h A(j, l)$ 表示图像 x_j 中出现频率最高的前面 h 个模式统计和， $\sum_{k=1}^q A(j, k)$ 表示图像 x_j 中出现的所有模式的频数统计和，式 (4) 是要找到出现频率比率超过阈值 n ($0 < n \leq 1$) 的最小模式子集。

鉴别能力： 理论上同类的图像应该具有相同的主导模式子集，但是由于受照明、噪声等其它因素的干扰，式 (4) 找到的模式子集在每一张图像中可能会出现差异，而如果能找到每一类图像集中具有一致性的模式子集，则该模式子集除了具有鲁棒性外，同时它们代表了该类图像的本质特性，在区分不同图像类时具有较强的鉴别力。为了找到能够

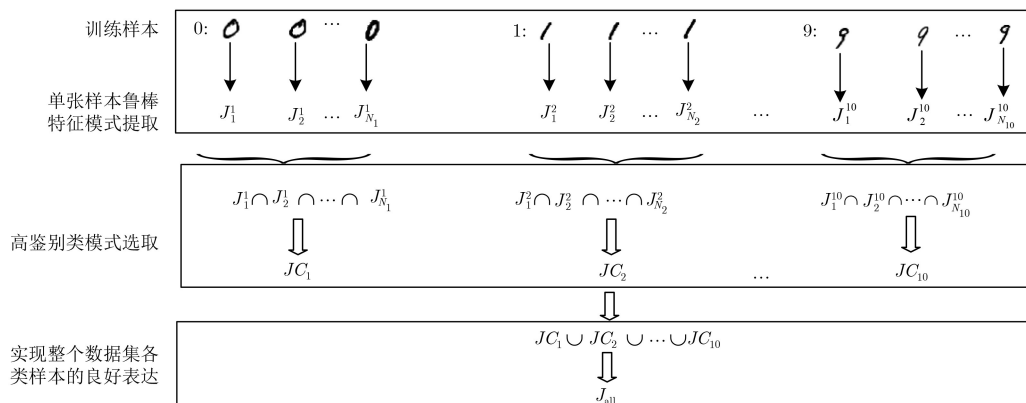


图3 主导模式学习模型

区分不同类别的高鉴别鲁棒主导模式，选用属于同一类的所有训练图像的主要模式的交集作为该类的判别主导模式集 JC_i 。

表示能力： 为了表示不同类别图像样本，最后采用所有类的判别主导模式集的并集来描述全部图像，并将其记为 J_{all} ， J_{all} 涵盖了所有类别的判别信息，因此具有较强的图像描述能力。

2.2.2 主导模式学习算法

输入： 每个训练图像 x_j 的ZMPB模式值，阈值参数 n

输出： 全局主导模式集 J_{all}

步骤1 统计训练图像 x_j 的每个ZMPB模式值出现的次数并按降序排列，结果记为 A 对 $h = 1, 2, \dots, q$ ，其中 q 为ZMPB模式类型的数量，判断是否满足 $\sum_{l=1}^h \frac{A(j, l)}{\sum_{k=1}^q A(j, k)} \geq n$ ，如果条件满足则选这 h 个模式作为训练图像 x_j 的主要模式，记为 $J_j = \{A[j, 1], A[j, 2], \dots, A[j, h]\}$ ；

步骤2 根据步骤1获取第 i 类所有训练图像的主要模式 $J_1^i, J_2^i, \dots, J_{N_i}^i$ ，取其交集获得判别主导模式集，记为 JC_i ，其中 $JC_i = J_1^i \cap J_2^i \cap \dots \cap J_{N_i}^i$ ；

步骤3 根据步骤2获取 C 类样本的判别主导模式集 JC_1, JC_2, \dots, JC_C ，取其并集获得全局主导模式集记为 J_{all} ，其中 $J_{all} = JC_1 \cup JC_2 \cup \dots \cup JC_C$ 。

2.3 空间池化与多尺度ZMPB模式特征提取

为了利用学到的主导模式来对图像进行描述，需要将图像分成格状子区域，然后通过空间池化操作实现对图像的整体描述。空间池化原理如图4所示，本文按照半重叠方式对图像进行分块，然后以块为基本单元对ZMPB的主导模式进行统计，获得相应直方图，并在块内将直方图特征进行归一化，最后将归一化后的每个块的特征组装到一起作为图像的特征表示。

为了实现在更大观测尺度描述图像视觉结构，本文可以在更大的局部区域中提取ZMPB模式。前文在 3×3 邻域中提取ZMPB模式是以单个像素点为基本单元，如果把其看作 1×1 尺度来处理，则还可以按 $2 \times 2, 3 \times 3, \dots$ 分块为基本单元来提取更大尺度空间的视觉模式结构。不妨以 2×2 分块尺度为例(图5)，在一个 6×6 的邻域中(图5(a))，按照 2×2 分块模式，求取块均值，并以均值表示该块，将 6×6 的邻域重新转换为 3×3 邻域(图5(b))，然后在转换后的 3×3 邻域中按照文所述的ZMPB模式提取方法计算 2×2 尺度上的ZMPB模式，其计算过程示意图见图5。如果将不同尺度的ZMPB模式特征进行组合，即可实现多尺度的ZMPB模式特征提取。

本文在实验中采用的图像识别算法采用了 1×1 和 2×2 两种分块尺度，在分别获取两种尺度

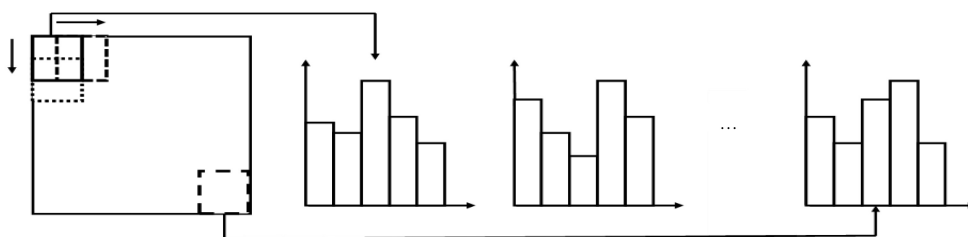


图4 空间池化示意图

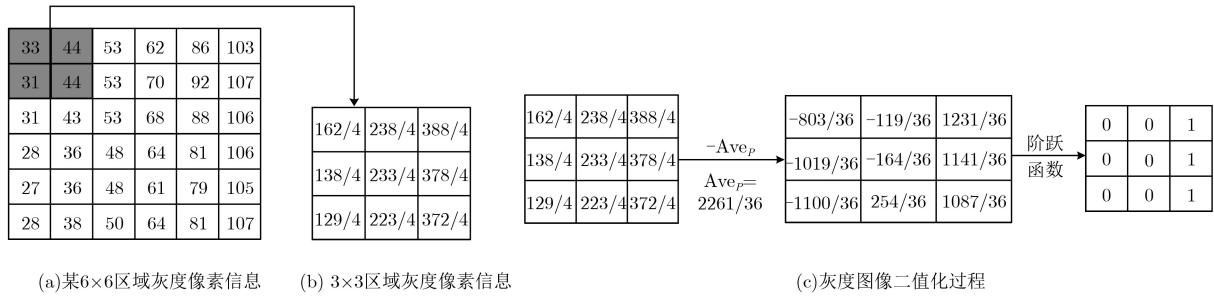


图5 2×2分块下ZMPB模式值计算示例图

下的ZMPB模式后，通过学习模型分别学习其主导模式子集，再通过空间池化操作得到图像的统计直方图特征，最后将1×1, 2×2两种尺度得到特征串联拼接在一起运用SVM分类器进行分类。

3 实验结果

为了评估算法性能，实验在ORL和YALE两个人脸公开数据集，MNIST手写数字公开数据集，以及本文搜集的非公开车标数据集进行了测试。本文所有的实验都是在Windows 7，编程环境为Matlab 2014，CPU的主频为3.70 GHz，内存为8 GB条件下进行的，且1×1尺度下阈值n设置为0.8，2×2尺度下阈值n设置为0.9。识别实验所用的分类器为Matlab自带的SVM分类工具箱。

3.1 人脸识别实验

ORL和YALE是两个公开的被广泛使用和测试的人脸数据库。ORL数据库包含40个个体，每个个体有10张照片，本文随机选取其中5张图片作为训练样本，其余5张作为测试样本，这样训练样本和测试样本均为200张图片。YALE数据库包含15个个体，每个个体采集了11张照片，本文随机选其中5张图像作为训练样本，剩下的6张作为测试样本，这样训练样本数为75，测试样本数为90。为了避免单次样本划分的偏向性，对ORL和YALE数据库都做了10次随机划分进行实验，最终识别率是10次测试结果的平均值。实验将YALE和ORL数据库中的所有的图片均归一化为40×40，以8×8半重叠方式做格状池化划分，然后对每一块分别以1×1, 2×2两种尺度计算ZMPB模式值，并分块统计两种尺度下的ZMPB模式的直方图特征，最后将1×1, 2×2两种尺度的特征串联拼接后再按块组合(从左往右，从上到下)到一起作为整幅图像的特征表示。

表1给出了LBP, GIMMRP和本文算法与其它已报导方法的检测结果，其它方法包括HOG^[16]算法, ACNN^[12]卷积神经网络方法, EMKFDA^[15]基于流形学习的方法, SHAO等人^[17]的度量学习方法

等。从表1中可以发现，本文提出的新算法在ORL数据集和YALE上的识别率优于其它已有算法，比原来最好的算法性能有较大的提升，例如在ORL数据集达到99.40%，比原来最好的SHAO和GIMMRP的97.5%识别率提高了1.9%，在YALE数据集上SHAO的方法可以做到97.5%，而本文方法达到99.30%，准确率提高了1.8%。

表1 各种算法人脸识别率比较(%)

识别算法	ORL	YALE
LBP	96.00	92.96
ART ^[10]	87.70	85.20
SRMKVS ^[11]	96.18	93.79
ACNN ^[12]	95.00	-
ALDR ^[13]	96.50	-
DTSA ^[14]	96.68	71.09
EMKFDA ^[15]	97.50	78.44
HOG ^[16]	96.00	93.67
SHAO等人方法 ^[17]	97.50	97.50
GIMMRP ^[9]	97.50	94.11
本文算法	99.40	99.30

3.2 手写数字识别实验

实验选取了MNIST手写数字公开数据库。其中MNIST图像大小为28×28，本文方法仍分别以1×1, 2×2两种尺度对图像进行分块计算其ZMPB模式值，并进行类似的8×8半重叠池化分块进行特征统计。实验主要比较了LBP, GIMMRP, 本文算法，以及近几年报导的其它方法的识别结果。从表2中可以看出，基于卷积神经网络的MCDNN效果最好，达到99.77%的识别准确率，本文算法次于MCDNN^[22]，但优于GIMMRP算法, LBP, HOG^[23]和数据驱动的其他几种方法，达到99.01%。

3.3 车标识别实验

车标图像数据集是在自然场景下获取的车辆图像基础上进行裁剪得到的，共收集了20类车标，其中每类包含100张图片。车标图片大小为40×40，

表2 各种算法在手写数字库MNIST的识别率比较(%)

识别算法	MNIST
LBP	93.56
CKELM ^[18]	96.80
MPDA ^[19]	89.91
WU等人方法 ^[20]	87.64
LIB-LLSVM+C-OCC ^[21]	98.39
MCDNN ^[22]	99.77
HOG ^[23]	97.25
GIMMRP ^[9]	98.91
本文算法	99.01

同样按照 1×1 , 2×2 两种尺度对图像进行分块统计ZMPB模式值,池化分块还是 8×8 半重叠模式。

由于是非公开数据集,表3只给出了LBP, GIMMRP和本文算法在不同训练样本比例下的识别结果,实验随机抽样了10次,最终识别率为10次的平均值。

由表3可知,在训练样本较少的情况下,本文算法已达到很高的识别率,达到99.87%,而且每次的检测准确率均优于GIMMRP算法和LBP方法。

表3 各种算法的车标识别率比较(%)

	训练样本数				
	10	20	30	40	50
LBP	97.95	99.42	99.69	99.87	99.92
GIMMRP ^[9]	99.64	99.88	99.95	99.96	99.96
本文算法	99.87	99.98	100	100	100

在样本比例占到30%时,算法准确率可以达到100%。实验结果表明,LBP在一定程度上受到训练样本数量的影响较大,识别率波动幅度远超过GIMMRP方法和本文算法。由于车标图像来源于自然场景中的图像,这从一定程度上验证了本文算法对光照变化、图像模糊等复杂背景下的车标图像具有很强的鉴别能力和鲁棒性。

3.4 本文算法与其它算法的性能比较

本文算法和LBP算法关系密切,同时又是对GIMMRP方法的改进,特别是引入了多尺度模式编码和模式学习技术,因此有必要对本文算法和LBP及GIMMRP算法性能进行进一步比较和分析。结果如表4所示。

从表4中不难发现,本文算法在单尺度条件下

表4 本文算法与相关算法性能比较

数据库		1×1识别率(%)	2×2识别率(%)	1×1+2×2识别率(%)	特征维度	1×1尺度单张图片特征提取时间(s)
YALE	本文算法	95.56	95.40	99.30	4050/5670/9720	0.020
	LBP	92.96			4779	0.016
	GIMMRP	94.11			10611	0.062
ORL	本文算法	97.70	97.45	99.40	7290/6966/14256	0.020
	LBP	96.00			4779	0.016
	GIMMRP	97.50			10611	0.061
车标	本文算法	99.11	99.10	99.76	4212/5670/9882	0.018
	LBP	97.95			4779	0.012
	GIMMRP	99.64			10611	0.053
MNIST	本文算法	98.32	98.93	99.01	720/792/1512	0.016
	LBP	93.56			2124	0.015
	GIMMRP	98.91			4716	0.044

识别率接近甚至超过GIMMRP算法,且优于LBP算法。而多尺度ZMPB特征明显优于单尺度ZMPB特征,同时在耗时上,单张图片ZMPB特征提取所耗时长一般稍高于LBP算法,但是远低于GIMMRP算法,平均耗时比GIMMRP算法约减少3倍,这是由于本文算法避免了GIMMRP算法中模式匹配的过程,改进了特征提取的效率,且本文算法加入了主导模式学习,在保证性能的前提下,大大降低了特征维度,加快了算法执行速度。例如,在YALE

实验中, 1×1 , 2×2 和 $1 \times 1 + 2 \times 2$ 尺度下的特征维度分别为4050, 5670, 9720,然而GIMMRP方法特征维度为10611。此外,对于LBP和本文方法,在 1×1 尺度下提取单张图片特征仅需0.016 s和0.020 s,而GIMMRP所需时间为0.062 s。

4 结束语

本文提出了一种基于零均值化的微观结模构式二值化编码ZMPB方法,且能够通过学习获取最适应数据集的主导二值模式特征子集,在保证性能的

前提下,大大降低特征维数,加快算法执行速度,而且本文方法在结合多尺度分块基础上,还可以进一步提升特征的鉴别能力。尽管算法在实验中获得了满意的结果,但是仍存在有待进一步深入探讨研究的工作。本文方法和传统的LBP方法在本质上应该是可以统一的,如何将它们归入到一个统一的框架下进行研究是一个值得深入思考的问题。另外可以考虑研究具有旋转不变性的ZMPB特征模式编码,以便增强算法的适用领域。

参 考 文 献

- [1] LOWE D G. Distinctive image features from scale-invariant key points[J]. *International Journal of Computer Vision*, 2004, 60(2): 91–110. doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94).
- [2] BAY H and TUYTELAARS T. SURF: Speeded up robust features[J]. *Computer Vision & Image Understanding*, 2006, 110(3): 404–417. doi: [10.1007/11744023_32](https://doi.org/10.1007/11744023_32).
- [3] MIKOLAJCZYK K and SCHMID C. A performance evaluation of local descriptors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1615–1630. doi: [10.1109/TPAMI.2005.188](https://doi.org/10.1109/TPAMI.2005.188).
- [4] ENGIN T, LEPETIT V, and FUA P. Daisy: An efficient dense descriptor applied to wide-baseline stereo[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(5): 815–830. doi: [10.1109/TPAMI.2009.77](https://doi.org/10.1109/TPAMI.2009.77).
- [5] DALAL N and TRIGGS B. Histograms of oriented gradients for human detection[C]. *IEEE Computer Society Conference on Computer Vision & Pattern Recognition*, San Francisco, USA, 2005: 886–893. doi: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177).
- [6] OJALA T, VALKEALAHTI K, OJA E, *et al.* Texture discrimination with multidimensional distributions of signed gray-level differences[J]. *Pattern Recognition*, 2001, 34(3): 727–739. doi: [10.1016/S0031-3203\(00\)00010-8](https://doi.org/10.1016/S0031-3203(00)00010-8).
- [7] OJALA T, PIETIKAINEN M, and MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971–987. doi: [10.1109/TPAMI.2002.1017623](https://doi.org/10.1109/TPAMI.2002.1017623).
- [8] LIAO S, LAW M W K, and CHUNG A C S. Dominant local binary patterns for texture classification[J]. *IEEE Transactions on Image Processing*, 2009, 18(5): 1107–1118. doi: [10.1109/TIP.2009.2015682](https://doi.org/10.1109/TIP.2009.2015682).
- [9] 张东波, 陈治强, 易良玲, 等. 图像微观结构的二值化表示与目标识别应用[J]. *电子与信息学报*, 2018, 40(3): 633–640. doi: [10.11999/JEIT170513](https://doi.org/10.11999/JEIT170513).
- [10] ZHANG Dongbo, CHEN Zhiqiang, YI Liangling, *et al.* Binarization representation of image microstructure and the application of object recognition[J]. *Journal of Electronics & Information Technology*, 2018, 40(3): 633–640. doi: [10.11999/JEIT170513](https://doi.org/10.11999/JEIT170513).
- [11] HAMDAN B and MOKHTAR K. Face recognition using Angular Radial Transform[J]. *Journal of King Saud University-Computer and Information Sciences*, 2016, 30(2): 141–151. doi: [10.1016/j.jksuci.2016.10.006](https://doi.org/10.1016/j.jksuci.2016.10.006).
- [12] ZHU Ningbo, TANG Ting, and TANG Shi. A sparse representation method based on kernel and virtual samples for face recognition[J]. *Optik-International Journal for Light and Electron Optics*, 2013, 124(23): 6236–6241. doi: [10.1016/j.ijleo.2013.05.017](https://doi.org/10.1016/j.ijleo.2013.05.017).
- [13] ZHANG Yuanyuan and ZHAO Dong. Adaptive convolutional neural network and its application in face recognition[J]. *Neural Processing Letters*, 2016, 43(2): 389–399. doi: [10.1007/s11063-015-9420-y](https://doi.org/10.1007/s11063-015-9420-y).
- [14] HUANG Pu and LAI Zhihui. Adaptive linear discriminant regression classification for face recognition[J]. *Digital Signal Processing*, 2016, 55: 78–84. doi: [10.1016/j.dsp.2016.05.001](https://doi.org/10.1016/j.dsp.2016.05.001).
- [15] WANG Sujing and ZHOU Chunguang. Face recognition using second-order discriminant tensor subspace analysis[J]. *Neurocomputing*, 2011, 74(12/13): 2142–2156. doi: [10.1016/j.neucom.2011.01.024](https://doi.org/10.1016/j.neucom.2011.01.024).
- [16] WANG Guoqiang and SHI Nianfeng. Embedded manifold-based kernel fisher discriminant analysis for face recognition[J]. *Neural Processing Letters*, 2016, 43(1): 1–16. doi: [10.1007/s11063-014-9398-x](https://doi.org/10.1007/s11063-014-9398-x).
- [17] SINGH G and CHHABRA I. Integrating global zernike and local discriminative HOG features for face recognition[J]. *International Journal of Image & Graphics*, 2016, 16(4): 1650021–1650042. doi: [10.1142/S0219467816500212](https://doi.org/10.1142/S0219467816500212).
- [18] SHAO Hong and CHEN Shuang. Face recognition based on subset selection via metric learning on manifold[J]. *Frontiers of Information Technology & Electronic Engineering*, 2015, 16(12): 1046–1058. doi: [10.1631/FITEE.1500085](https://doi.org/10.1631/FITEE.1500085).
- [19] DING Shifei and GUO Lili. Extreme learning machine with kernel model based on deep learning[J]. *Neural Computing & Applications*, 2017, 28(8): 1975–1984. doi: [10.1007/s00521-015-2170-y](https://doi.org/10.1007/s00521-015-2170-y).
- [20] ZHOU Yang and SUN Shiliang. Manifold partition discriminant analysis[J]. *IEEE Transactions on Cybernetics*, 2017, 47(4): 830–840. doi: [10.1109/TCYB.2016.2529299](https://doi.org/10.1109/TCYB.2016.2529299).

- [20] WU Tingfang and LIN C J. Probability estimates for multi-class classification by pairwise coupling[J]. *Journal of Machine Learning Research*, 2004, 5(4): 975–1005.
- [21] SCHMIDHUBER J, CIRES D, and MEIER U. Multi-column deep neural networks for image classification[C]. *IEEE Conference on Computer Vision & Pattern Recognition*, Rod Aproveendis, USA, 2012: 3642–3649. doi: [10.1109/CVPR.2012.6248110](https://doi.org/10.1109/CVPR.2012.6248110).
- [22] ZHANG Ziming and LADICKY L. Learning anchor planes for Classification[C]. *Advances in Neural Information Processing Systems*, Granada, Spain, 2011: 1611–1619.
- [23] EBRAHIMZADEH R and JAMPOUR M. Efficient handwritten digit recognition based on histogram of oriented gradients and SVM[J]. *Annals of the Rheumatic Diseases*, 2014, 104(9): 10–13. doi: [10.5120/18229-9167](https://doi.org/10.5120/18229-9167).
- 张东波: 男, 1973年生, 博士, 教授, 研究方向为计算机视觉、模式识别.
- 易良玲: 女, 1993年生, 硕士, 研究方向为计算机视觉、机器学习.
- 许海霞: 女, 1979年生, 博士, 副教授, 研究方向为机器视觉、模式识别.
- 张莹: 男, 1972年生, 博士, 副教授, 研究方向为机器人控制、模式识别、高维可视化处理.