

## 圆形图像抗旋转高效高鉴别特征表示方法<sup>\*</sup>

张东波<sup>1,2</sup>, 陈红磊<sup>1,2</sup>, 文登伟<sup>1,2</sup>, 汤红忠<sup>1,2</sup>, 许海霞<sup>1,2</sup>



<sup>1</sup>(湘潭大学 信息工程学院, 湖南 湘潭 411105)

<sup>2</sup>(机器人视觉感知与控制技术国家工程实验室(湖南大学), 湖南 长沙 410012)

通讯作者: 张东波, E-mail: zhadonbo@163.com

**摘要:** 根据圆形图像目标中心对称特性, 提出一种基于空间位置描述的抗旋转高效高鉴别二值模式特征提取方法. 该方法在特征计算时, 通过径向变换重新构建局部坐标系, 并在此基础上实现具有抗旋转变换的空间对称区域的局部二值模式提取; 同时, 池化操作采用了具有旋转不变性的环状空间划分, 从而保证了最终特征描述的抗旋转变换能力. 该方法分别在欧元硬币数据集、QQ表情数据集、车标数据集上进行了测试, 识别准确率最高分别达到 100%, 100%, 97.07%. 在具有旋转情况的欧元硬币和 QQ 表情数据集中, 优于传统的 LBP 特征和 HOG 特征, 而且算法具有高效性, 其单点计算时间只需 0.045ms.

**关键词:** 特征提取; 抗旋转; 图像识别; 二值模式

**中图法分类号:** TP311

中文引用格式: 张东波, 陈红磊, 文登伟, 汤红忠, 许海霞. 圆形图像抗旋转高效高鉴别特征表示方法. 软件学报, 2019, 30(9): 2904–2917. <http://www.jos.org.cn/1000-9825/5566.htm>

英文引用格式: Zhang DB, Chen HL, Wen DW, Tang HZ, Xu HX. Anti-rotation and efficient discriminative feature representation method for circular images. Ruan Jian Xue Bao/Journal of Software, 2019, 30(9): 2904–2917 (in Chinese). <http://www.jos.org.cn/1000-9825/5566.htm>

## Anti-rotation and Efficient Discriminative Feature Representation Method for Circular Images

ZHANG Dong-Bo<sup>1,2</sup>, CHEN Hong-Lei<sup>1,2</sup>, WEN Deng-Wei<sup>1,2</sup>, TANG Hong-Zhong<sup>1,2</sup>, XU Hai-Xia<sup>1,2</sup>

<sup>1</sup>(College of Information Engineering, Xiangtan University, Xiangtan 411105, China)

<sup>2</sup>(Robot Visual Perception & Control Technology National Engineering Laboratory, Changsha 410012, China)

**Abstract:** According to the center symmetric characteristics of circular image objects, this study proposes an anti-rotation and efficient discriminative binary feature extraction method based on pairs of spatial symmetry structure. This method reconstructs the local coordinate system by radial transform during feature computation, based on it, then local binary pattern with anti-rotation of spatial symmetry regions are extracted. Meanwhile, the annular space is adopted to achieve rotation invariability during the feature pooling operation, which ensures the anti-rotation ability of final feature description. This method are tested in the euro coins, QQ expression, and car logo data set, and the recognition accuracy reached 100%, 100%, and 97.07% respectively, which is superior to traditional LBP and HOG features in euro coins and QQ expression datasets. Moreover, the algorithm is efficient, and the computation time for single point feature extraction is only 0.045 ms.

**Key words:** feature extraction; anti-rotation; image recognition; binary pattern

\* 基金项目: 国家自然科学基金(61602397, 61573299); 湖南省自然科学基金(2017JJ2251, 2017JJ3315); 湖南省重点学科建设项目

Foundation item: National Natural Science Foundation of China (61602397, 61573299); Natural Science Foundation of Hunan Province (2017JJ2251, 2017JJ3315); Key Discipline Construction Project of Hunan Province

收稿时间: 2017-05-24; 修改时间: 2017-10-23; 采用时间: 2018-02-06; jos 在线出版时间: 2018-04-27

CNKI 网络优先出版: 2018-04-27 14:58:06, <http://kns.cnki.net/kcms/detail/11.2560.TP.20180427.1457.005.html>

现实生活中存在大量的圆形图像目标识别问题,例如车标、图标、图徽、商标、标签、硬币、装饰物上的圆形花纹和图案等.在很多场合,圆形图像印制、摆放的角度具有随意性,因此图案常出现不同角度的旋转.为了解决现实场景中圆形图像目标的识别,首先要保证在特征提取和图像描述时的抗旋转变换能力,这样才能保证在只获得少量特定角度训练样本前提下实现任意角度圆形图像目标的识别,从而避免搜集和存储大量的各种角度的圆形图像样本.以往的文献中没有专门针对圆形图像目标识别的研究,但是在各种应用案例中有相关的研究报告,例如对硬币图案的识别、车标的识别、LOGO 标签的识别等.

图像目标识别一直是计算机视觉领域的热点研究问题,其通常包括图像预处理、特征提取与描述、分类器设计与学习等环节,其中,特征提取与描述最为关键,需要综合考虑鉴别能力、计算复杂性和鲁棒性等因素.针对圆形图像目标的识别问题,我们希望构建一种具有广泛适用性的抗旋转高效高鉴别的特征提取与描述方法.

图像目标检测与识别的常用特征描述算子分为浮点型和二值型两类:浮点型描述子通常基于梯度计算,通过梯度值和梯度方向信息进行描述,著名的有 SIFT<sup>[1]</sup>,SURF<sup>[2]</sup>,MROGH<sup>[3]</sup>,HOG<sup>[4]</sup>等,其优点是鉴别能力强,但计算量大比较耗时.二值描述算子中有代表性的是 LBP<sup>[5]</sup>,BRIEF<sup>[6]</sup>,ORB<sup>[7]</sup>,BRISK<sup>[8]</sup>,FREAK<sup>[9]</sup>等,二值描述算子通常直接比较图像像素之间灰度大小,优点在于计算简单,计算和匹配效率高,缺点是由于量化信息损失,其鉴别能力较弱.

为了实现抗旋转特征提取和图像描述,需要考虑两个方面的问题:一是在特征计算时应该具有抗旋转变换能力,二是通过池化操作对图像做整体描述时应该具有旋转不变性.方向校正是特征计算和池化操作常用的抗旋转技术手段,SIFT 和 SURF 算子通过估计局部主梯度方向作为参考方向,然后做适当的方向校准实现抗旋转变换.但由于梯度方向估计可能存在 $\pm 20^\circ$ 的误差<sup>[10]</sup>,因此易导致偏差.LBP 为了获得旋转不变特征模式<sup>[5]</sup>,则通过排序策略找到具有最小模式值的模式特征作为代表模式.该处理策略在梯度方向直方图中也可以借用,例如找到最大梯度方向,然后以此为基准,按顺时针或逆时针方向重新排列梯度方向直方图.但是此类方法易受噪声和其他因素干扰,有时难以获得稳定的排序关系.也有研究人员采用成对特征点的梯度方向夹角取代方向,由于夹角和旋转无关,因此可实现旋转不变描述<sup>[11]</sup>.构建局部坐标系也是在特征计算阶段常用的一种抗旋转变换策略,例如樊彬的局部旋转不变坐标系<sup>[10]</sup>、文献[12]采用的径向梯度变换,局部坐标系在获得旋转不变性能力上,具有严谨的理论支撑,有相对稳定的性能.

在对图像整体描述时的特征池化操作中,通常会采用格状空间划分,如果不做方向校正的话,格状空间不具备旋转不变性.现有文献中,最典型的旋转不变描述有环状空间划分<sup>[13]</sup>和基于灰度排序的池化策略<sup>[14]</sup>.

- 环状空间划分原理简单,具有天然的旋转不变性,因此在做旋转不变描述时,其特征表达最为稳定.但由于对图像空间的描述比较粗糙,造成空间信息损失,因此特征鉴别能力不强.这也是很多特征描述方法不采用环状空间划分而往往采用格状空间划分的原因,例如 SIFT,HOG 等;
- 灰度排序则通过对灰度区间的量化和排序实现和旋转无关的图像描述,其在实际应用中有不错的性能表现.但是灰度强度易受照明的影响,此外,该方法对目标遮挡和污渍敏感,因此实际应用中也有较大的局限性.

鉴于上述分析,本文提出一种针对圆形图像目标的抗旋转高效高鉴别特征提取与描述方法.该方法为了解决特征计算的抗旋转能力,采用了径向变换技术,同时,充分利用了圆形图像目标本身呈中心对称的外观特点,采用环状空间特征池化策略,并且利用空间对称区域结构的描述提升了环状空间描述的鉴别能力.此外,由于该方法属于二值模式特征,其继承了传统二值模式特征计算和匹配高效的优点.

## 1 抗旋转对称结构模式特征提取

### 1.1 现有LBP模式计算原理及缺陷

局部二值模式(local binary pattern,简称 LBP)是一种简单有效的特征提取算法,原  $LBP_{p,r}$  是通过比较中心像素和半径为  $r$  的圆形邻域的  $P$  个像素点的灰度大小来实现微观邻域结构的描述.LBP 具有原理简单、单调照明变化不变性等优点,改进后的 LBP 具有旋转不变性,因此广泛应用在图像匹配与图像识别等领域,特别是在纹理

分析与人脸识别领域.

为了降低模式数目,可以只保留有重要意义的模式结构,因此,研究人员提出了统一模式  $LBP_{p,r}^{u_2}$ ,统一模式在一次二进制循环计算中只产生两次变化.为了实现旋转不变性,Ojala 等人<sup>[5]</sup>又提出了旋转不变 LBP,通过旋转圆形邻域可以得到一系列初始定义 LBP 值,取其中最小值作为该邻域 LBP 值,旋转不变模式记作  $LBP_{p,r}^i$ .结合统一和旋转不变模式原理,可以得到旋转不变统一模式  $LBP_{p,r}^{iu_2}$ .以邻域  $p=8$  为例,其模式数目从最初的  $2^8=256$  种减少到 10 种.由于兼具统一模式和旋转不变模式的特点,  $LBP_{p,r}^{iu_2}$  在图像识别领域得到广泛应用,特别是在需要提取旋转不变特征の場合.旋转不变统一二值模式由下式(1)计算:

$$LBP_{p,r}^{iu_2} = \begin{cases} \sum_{p=0}^{p-1} S(g_p - g_c), & U(LBP_{p,r}) \leq 2 \\ p + 1, & \text{otherwise} \end{cases} \quad (1)$$

其中,  $U(LBP_{p,r}) = |S(g_{p-1} - g_c) - S(g_0 - g_c)| + \sum_{p=0}^{p-1} |S(g_p - g_c) - S(g_{p-1} - g_c)|$ ,  $S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ ,  $g_c$  表示中心点像素值,  $g_i$  表示第  $i$  个邻域点像素值.  $S$  是符号函数,  $U(LBP_{p,r}) \leq 2$  计算得到的是统一不变模式.图 1 是  $LBP_{p,r}^{iu_2}$  计算示例.

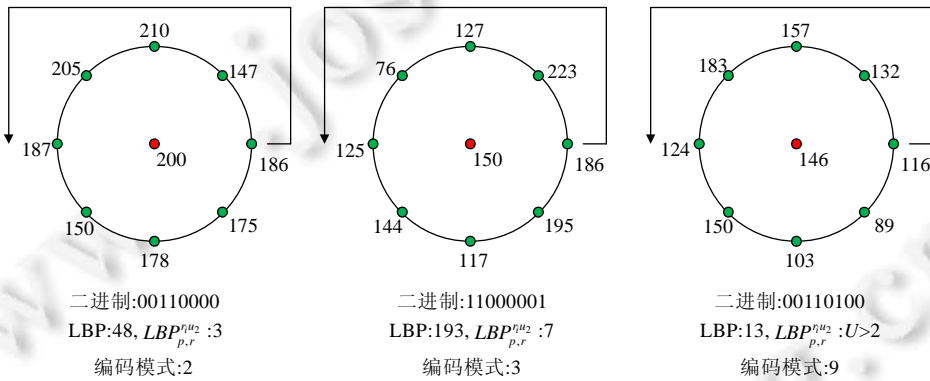


Fig.1 Illustration of the calculation of LBP and  $LBP_{p,r}^{iu_2}$

图 1 LBP 和  $LBP_{p,r}^{iu_2}$  计算示例图

考虑到圆形图案的外观特点,在进行图像整体描述时,为了获得旋转不变性,我们很自然地考虑采用环形特征池化操作.但是环形空间划分本身鉴别能力弱,不利于后续识别.如果采用经典的旋转不变统一模式特征  $LBP_{p,r}^{iu_2}$ ,其只能描述 10 种结构模式,不利于图像细节的描述,会影响特征的鉴别能力.因此,为了既能保留环形空间在处理旋转不变描述时的天然优势,同时又能弥补环形空间划分造成的空间信息损失,有必要利用更多的空间结构信息来有效提高特征的鉴别能力.本文充分利用圆形图案中心对称的特点,提出一种基于空间对称位置描述的二值模式特征提取方法.

### 1.2 基于局部径向坐标变换的旋转不变二值模式及其近似计算

为了获得旋转不变描述,我们引入局部径向坐标变换技术,并按照给定样本点的径向与切向方向选取邻域点进行编码.提取样本点  $b$  的旋转不变二值模式原理如图 2 所示:  $c$  为图像中心,沿  $cb$  方向称为径向坐标轴,垂直  $cb$  方向称为切向坐标轴,  $b$  为新的局部坐标系原点.分别在径向  $r$  和切向  $t$  方向找到对称的与  $p$  点的距离为  $d$  的 4 个邻域点,沿逆时针分别记作  $b_1, b_2, b_3, b_4$ ,其中,径向方向上的邻近点为  $b_1, b_3$ ,切向方向上的邻近点为  $b_2, b_4$ ,然后按照  $b_1, b_2, b_3, b_4$  的顺序与  $b$  点的像素值进行比较,由公式(2)得到二进制数编码  $T=(T_1T_2T_3T_4)$ :

$$T_i = \begin{cases} 1, & I(b_i) > I(b) \\ 0, & I(b_i) \leq I(b) \end{cases}, i=1,2,3,4 \quad (2)$$

将 4 位二进制数  $T$  按下式转化为 10 进制数,即可得到对应的编码值:

$$F=8 \times T_1+4 \times T_2+2 \times T_3+T_4 \quad (3)$$

其中,  $I(b)$  为  $b$  点的像素值,  $I(b_i)$  为所取点  $b_i$  的像素值,  $F$  为编码值. 该二值编码方法可以描述 16 种二值模式, 与旋转不变统一模式相比, 其不仅具有旋转不变性, 而且可以表示更多的细节信息.

在构建局部坐标系统时, 需要进行浮点数插值计算, 这需要大量的计算时间代价. 为了避免插值计算, 提高计算速度, 我们可以采用近似计算, 其原理如图 3 所示: 在圆形区域内, 将图像中心径向方向以  $45^\circ$  为单位, 划分 8 个扇区, 如图 3 虚线所分隔扇区. 环状空间中采样点的近似编码计算示例如图 4~图 7 所示.

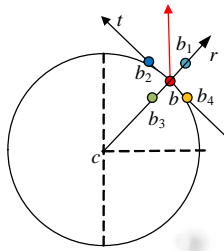


Fig.2 Diagram of local radial coordinate transformation  
图 2 局部径向坐标变换示意图

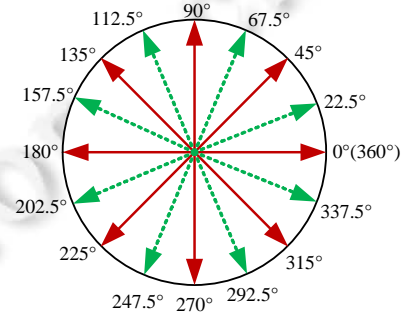


Fig.3 Sector schematic for radial approximate coding  
图 3 径向近似计算的扇区示意图

如图 4 所示, 假设点  $A, B, C$  分别为圆上一点, 它们与图像中心夹角分别为  $\theta_A < 22.5^\circ, 22.5^\circ < \theta_B < 67.5^\circ, 67.5^\circ < \theta_C < 112.5^\circ$ , 图 5~图 7 给出了 3 个点在  $3 \times 3$  邻域内的近似编码计算实例. 此时, 点  $A$  的编码起始位置  $b_1$  为邻域水平方向右边像素点, 即图 5 中像素值为 120 的位置, 按逆时针方向分别取右、下、左、上这 4 个点进行编码. 对于点  $B$  来说, 则编码起始位置  $b_1$  为邻域右上角像素点, 即图 6 中像素值为 137 的位置, 此时按右上、右下、左下、左上的顺序取对角线的 4 个点, 实现逆时针编码. 而点  $C$  的编码起始位置  $b_1$  则为邻域垂直方向上方像素点, 即图 7 中像素值为 128 的位置, 按上、右、下、左顺序取 4 个点实现逆时针编码. 点  $A, B, C$  编码计算过程示意图分别如图 5~图 7 所示.

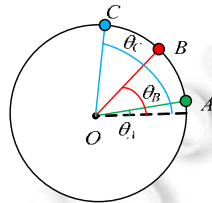


Fig.4 Location of points A, B, C  
图 4 点 A, B, C 所在位置图

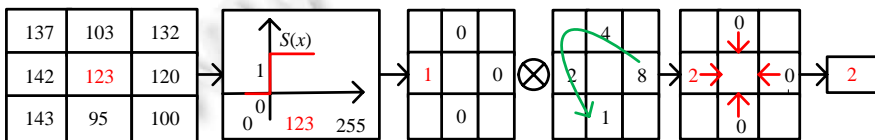


Fig.5 Encoding calculation process of point A  
图 5 点 A 编码计算过程图

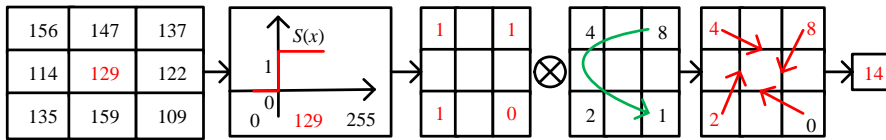


Fig.6 Encoding calculation process of point B

图 6 点 B 编码计算过程图

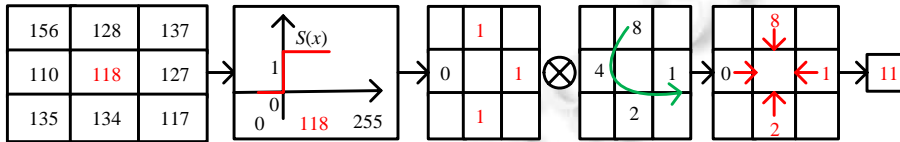


Fig.7 Encoding calculation process of point C

图 7 点 C 编码计算过程图

### 1.3 环形空间对称结构模式提取

在对圆形物体 ROI 区域中的每个点进行编码后,考虑到圆形图案中心对称的特点,我们采用环状空间特征池化策略;同时,为了提升环状空间描述的鉴别能力,我们提出空间对称结构对的方式提取成对点的特征模式。

由于圆形物体的 ROI 区域为圆形,因此可将该区域按照等面积的原则划分为若干个交叠的圆环区域,如图 8 所示,其中,第  $m$  个圆环包含的区域半径为  $r_m \in [r \times \sqrt{(m-1)}, r \times \sqrt{(m+1)}], m = 1, 2, \dots, n, r = R/\sqrt{n+1}$ , 其中,  $n$  为圆环的个数,  $R$  为圆形 ROI 区域半径.假设  $(x_0, y_0)$  为 ROI 圆形区域的中心点,  $(i, j)$  为区域内任意一点,  $(i', j')$  为  $(i, j)$  的对称点,则两者之间的坐标关系由公式(4)求得:

$$\begin{cases} i' = 2x_0 - i \\ j' = 2y_0 - j \end{cases} \quad (4)$$

分别对  $(i, j)$  与其对称点  $(i', j')$  按照第 1.2 节所述的径向坐标变换方式计算其局部二值模式和编码值,其中,点  $(i, j)$  编码值记为  $F_1$ , 点  $(i', j')$  编码值记为  $F_2$ . 为了确保点对模式编码与旋转无关,有必要按照大小排序关系进行组合.如果  $F_1 > F_2$ , 则点对模式表示为  $(F_1, F_2)$ , 否则表示为  $(F_2, F_1)$ . 统计圆环内所有点对, 按照出现频率即可获得该圆环内所有点对模式的直方图描述, 记作  $h_m$ . 根据组合关系, 模式对有  $16^2 = 256$  种组合, 因此,  $h_m$  是一个 256 维的特征向量. 考虑到对称性, 实际计算时, 点  $(i, j)$  只需要在圆环上半部分取样, 这可以避免不必要的重复统计.

影响模式对统计稳定性的因素主要有两个:一是 ROI 区域中心位置估算误差,二是图像噪声和干扰.为了获得稳定的统计结果,我们不是直接统计单个点对的直方图,而是在当前点  $(i, j)$  和对称点  $(i', j')$  一定邻域内各自寻找代表性模式作为该点的模式表示.根据实际图像不同的特点,我们设计了两种方式.

- 如果圆形图案纹理细节信息比较丰富,则我们采用主模式选择编码.即:在点  $(i, j)$  的  $M \times M$  邻域内统计出现次数最多的编码模式值,将该模式值作为点  $(i, j)$  的代表模式,记为  $F_1$ ;同样统计对称点  $(i', j')$  的代表模式,记为  $F_2$ ,我们将这种特征提取方式称为邻域代表模式;
- 如果圆形图案主要由线条、边缘轮廓构成,没有丰富的纹理细节,则我们采用主梯度选择编码.即:寻找点  $(i, j)$  的  $M \times M$  邻域内梯度幅值最大的点,并取其径向变换二值模式值作为点  $(i, j)$  的代表模式,记为  $F_1$ ;按同样原理确定对称点  $(i', j')$  的代表模式,记为  $F_2$ .我们将这种特征提取方式称为邻域主梯度代表模式.需要注意的是,在计算梯度时,同样应该在新的局部径向  $r, t$  坐标系统内进行计算,这样才能保证旋转不变性.

将圆形物体所在的 ROI 圆形区域按等面积划分为  $n$  个交叠圆环,计算每个圆环内的对称结构模式对;然后统计圆环内的特征模式对直方图,记为  $h_m$  (256 维的特征向量);最终,按从内环到外环的顺序组装起来得到最终的图像特征描述:

$$H = (h_1, h_2, \dots, h_n) \quad (5)$$

其中,  $h_m$  表示第  $m$  个圆环的特征向量,  $m=1, 2, \dots, n$ .

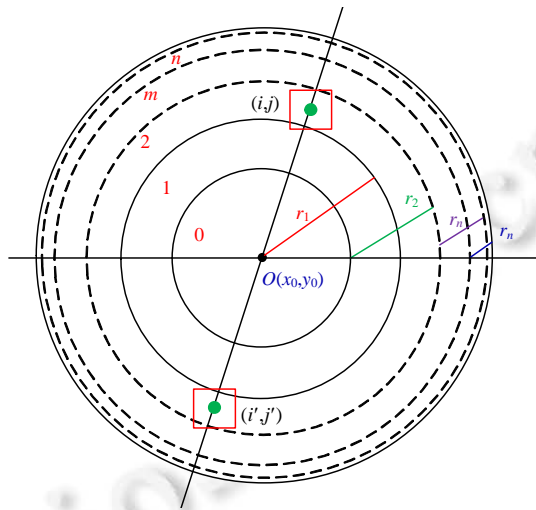


Fig.8 Schematic diagram of symmetrical structure

图 8 对称结构模式示意图

#### 1.4 抗旋转对称结构模式特征提取算法

抗旋转对称结构模式特征提取算法步骤如下。

- 1) 提取圆形图像 ROI 区域,并确定其中心位置;
- 2) 对于 ROI 区域内每个点,按照局部径向坐标变换选取 4 个邻域点,并采用近似计算方式,计算该点的抗旋转编码模式值;
- 3) 将整个 ROI 区域按照等面积的原则划分出相互交叠的  $n$  个圆环;
- 4) 依次统计第  $m$  个( $m=1, \dots, n$ )圆环内关于圆形区域中心点  $O$  对称的点对  $(i, j)$  与  $(i', j')$  的模式,并选择主模式编码或主梯度选择编码方式,统计完圆环内所有点对后,按照出现频率即可获得该圆环内所有模式对组合的直方图描述,记作  $h_m$ ;
- 5) 将  $n$  个圆环提取到的特征向量,按照从内环到外环顺序组装得到最终的图像特征描述,记为

$$H=(h_1, h_2, \dots, h_n).$$

## 2 实验结果及分析

### 2.1 实验数据集介绍

实验采用了欧元硬币、QQ 表情和车标这 3 个数据集。欧元硬币数据集由 12 个主要欧盟国家的欧元硬币组成,所有国家硬币正面图案相同,反面图案由每一个国家自行设计,均不相同。硬币数据集包含 2 欧元、1 欧元、50 欧分、20 欧分、10 欧分 5 种币值,我们采集了不同币值的正反面硬币图像 8 430 张。其中,2 欧元 2 603 张,1 欧元 2 240 张,50 欧分 1 030 张,20 欧分 1 291 张,10 欧分 1 266 张。图 9 给出了德国、卢森堡和葡萄牙的 5 种币值的反面图案。硬币检测的目的是通过识别反面图案实现硬币国别的判别。

QQ 表情数据集是截取腾讯聊天软件中的表情获得的,共包括 52 种不同的表情,每种表情图片按 5 度加以旋转,可扩充为 72 张图片,总共获得 3 744 张图片。QQ 表情图片示例如图 10 所示。

车标数据集通过对各种车型的网上实物图像截取、建立。我们专门选择了 15 种近似圆形的车标图案,其中,每种车型截取了 100 张图片,共 1 500 张,部分车标图片如图 11 所示。



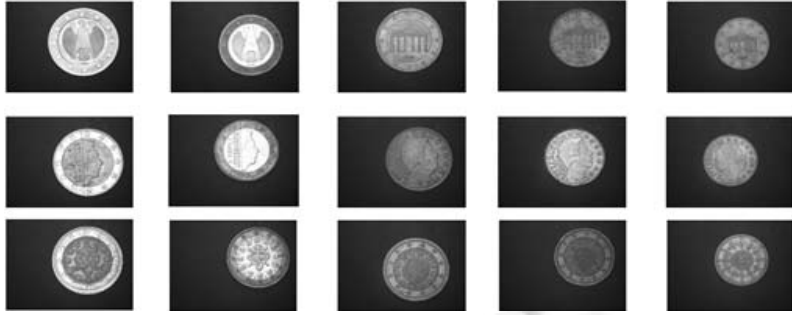


Fig.9 Part of the euro coin figures

图 9 部分欧元硬币图片



Fig.10 Part of QQ expression pictures

图 10 部分 QQ 表情图片



Fig.11 Part of the car logo pictures

图 11 部分车标图片

表 1 给出了 3 个数据集的基本信息.我们之所以采用上述 3 个数据集,是因为这 3 个数据集有各自不同的特性.它们在色彩、纹理、光照、图案复杂程度方面有明显差异,这可以充分检验方法的适应性和鲁棒性.

**Table 1** Experimental dataset basic information

表 1 实验数据集信息

数据集	类别数	样本数	图像类型	图像特点
欧元硬币	13	8 430	灰度	图案较复杂,纹理信息丰富,亮度差异变化大,有划痕、污渍和磨损等干扰
QQ表情	52	3 744	彩色	图案简单,因为色差表现出一定纹理变化
车标	12	1 500	彩色	图案较简单,自然场景下获取,背景、照明和视角均有变化,纹理变化少

数据图像中目标的 ROI 圆形区域对本文算法特征提取具有重要作用.由于欧元硬币图像有确定的单一背景,通过简单的背景差分和形态学操作,较容易确定硬币的位置和 ROI 区域.而对于 QQ 表情和车标图片,由于我们重点关注的是圆形图像识别的问题,因此在收集数据集时,我们已经人为地将圆形图像目标从原图中剪切并保存下来,后续算法只需在剪切下来的方形子图中做内切圆即可确定其 ROI 圆形区域.

2.2 实验结果与分析

本文方法是基于圆形图像呈现中心对称的特点而设计的,通过径向变换技术和邻域模式选择策略,特征表达对于旋转和噪声具有较强的鲁棒性.因此,该方法命名为“中心-对称鲁棒二值模式”,简记为 CS\_RBP(central symmetry robust binary pattern).由于本文方法和传统的LBP方法以及HOG方法在特征提取和采样方式上相似,因此我们主要比较LBP,HOG方法,以及采用类似的中心对称结构和邻域模式选取策略的LBP方法和本文提出的CS\_RBP方法在欧元硬币数据集、QQ表情数据集和车标数据集这3个圆形图案数据集上的实验结果.其中,分类器模型选择的是线性SVM.

2.2.1 关键参数选取实验

有两个重要参数对CS\_RBP方法检测性能影响较大:一是环形空间的数目;二是在选择邻域代表模式时,邻域窗口的大小.为了获得最佳识别性能,有必要通过实验方式确定合适的参数值.以硬币图像为例,表2给出了将硬币图像归一化为200×200后,邻域M×M取为3×3,5×5,7×7,9×9,11×11这5种模式,圆环个数m分别取16,18,20,22,24,26,28等情况下的硬币国别识别结果.

Table 2 M×M neighborhood size and number of rings m selection experiment

表2 M×M邻域大小及圆环个数m选取实验

M×M \ m	16	18	20	22	24	26	28
3×3	99.95% (8426/8430)	99.94% (8425/8430)	<b>99.96%</b> <b>(8427/8430)</b>	<b>99.96%</b> <b>(8427/8430)</b>	<b>99.96%</b> <b>(8427/8430)</b>	99.95% (8426/8430)	99.95% (8426/8430)
5×5	99.95% (8426/8430)	99.95% (8426/8430)	99.96% (8427/8430)	<u>99.98%</u> <u>(8428/8430)</u>	99.96% (8427/8430)	99.98% (8428/8430)	<b>99.99%</b> <b>(8429/8430)</b>
7×7	<u>99.96%</u> <u>(8427/8430)</u>	<u>99.96%</u> <u>(8427/8430)</u>	99.96% (8427/8430)	<u>99.96%</u> <u>(8427/8430)</u>	<u>99.98%</u> <u>(8428/8430)</u>	<b>100%</b> <b>(8430/8430)</b>	<u>99.99%</u> <u>(8429/8430)</u>
9×9	<b>99.94%</b> <b>(8425/8430)</b>	99.85% (8417/8430)	99.85% (8417/8430)	99.89% (8421/8430)	99.89% (8421/8430)	99.88% (8420/8430)	99.91% (8422/8430)
11×11	99.72% (8406/8430)	99.74% (8408/8430)	99.66% (8401/8430)	99.61% (8397/8430)	99.70% (8405/8430)	99.74% (8408/8430)	<b>99.75%</b> <b>(8409/8430)</b>

从实验结果观测可知,邻域大小的选取比环状空间数目更为重要.在同等邻域大小下,环状空间数目对识别性能的影响比较小,检测结果波动幅度很小.表2中,每一行加粗的数据是在该邻域大小下,在某环状数目下的最好识别结果.在3×3邻域时,环数为20,22,24时有最好结果;而在5×5邻域时,环数为28时有最好结果;7×7邻域时,环数为26时有最好结果;9×9邻域时,环数为16时有最好结果;11×11邻域时,则环数为28时有最好结果.在同等邻域下,环数取多少有最好结果没有明显的规律.同时可观测到:随着邻域增大,9×9和11×11两组实验结果的波动相对比邻域窗口较小的3×3,5×5和7×7的实验结果波动幅度要大一些.另外,观测每一列的数据可知:在环数固定后,邻域过大和过小都会导致识别性能下降.在所有7列数据中,除了环数22时,5×5邻域有最好结果外(下划线数据),其他列中均是在邻域7×7时取得最佳识别性能.最终的最好识别结果为100%,其中邻域为7×7、环数26.鉴于上述实验结果,7×7邻域、环数26是一组较佳的参数.在后续实验中,我们均选用该组参数进行实验测试和比较.

2.2.2 不同算法在3个数据集上的实验

针对欧元硬币数据集、QQ表情数据集和车标数据集,我们分别采用不同比例的训练样本进行实验.用来比较的LBP特征是最常用的LBP统一旋转不变模式 $LBP_{p,r}^{n/2}$ ,选取了(p,r)=(8,1)和(p,r)=(8,2)两组参数.在我们的CS\_RBP方法中,模式编码时选取了d=1和d=2两组参数.同时,考虑到本文方法改进时所采取的对称结构、邻域代表模式等手段同样可用于原来的LBP特征提取,因此, $LBP_{p,r}^{n/2}$ 也可以结合对称模式实现单点、基于邻域的代表模式或主梯度代表模式这3种算法形式,因此,我们总共比较了 $LBP_{8,1}^{n/2}$ 、 $LBP_{8,1}^{n/2}$ 单点、 $LBP_{8,1}^{n/2}$ 邻域代表模式、 $LBP_{8,1}^{n/2}$ 邻域主梯度代表模式、 $LBP_{8,2}^{n/2}$ 、 $LBP_{8,2}^{n/2}$ 单点、 $LBP_{8,2}^{n/2}$ 邻域代表模式、 $LBP_{8,2}^{n/2}$ 邻域主梯度代表模式、CS\_RBP<sub>4,1</sub>单点、CS\_RBP<sub>4,1</sub>邻域代表模式、CS\_RBP<sub>4,1</sub>邻域主梯度代表模式、CS\_RBP<sub>4,2</sub>单点、CS\_RBP<sub>4,2</sub>邻域代表模式、CS\_RBP<sub>4,2</sub>邻域主梯度代表模式等算法.此外,本文方法和实验所用的LBP类方法均属于密集采



样,与传统 HOG 特征方法的描述方式相似,因此,HOG 特征描述也是一种适合和本文特征方法对比的方法.表 3~表 5 给出了上述方法在不同训练集占比下对硬币、QQ 表情和车标进行识别的实验结果.为了在 7×7 邻域、环数 26 参数下取得好的结果,其中图片尺寸均归一化为 200×200.在灰度欧元硬币数据集上,  $LBP_{p,r}^{n_u}$  提取的特征维数为 260 维,  $LBP_{p,r}^{n_u}$  单点、  $LBP_{p,r}^{n_u}$  邻域代表模式、  $LBP_{p,r}^{n_u}$  邻域主梯度代表模式的特征维数 1 320 维,CS\_RBP 单点、CS\_RBP 邻域代表模式、CS\_RBP 梯度邻域主梯度代表模式的特征维数均为 3 536 维,在彩色 QQ 表情数据集和车标数据集上,分别提取 R,G,B 这 3 个通道特征,再合并为一个最终特征,此时  $LBP_{p,r}^{n_u}$  提取的特征维数为 780 维、  $LBP_{p,r}^{n_u}$  单点、  $LBP_{p,r}^{n_u}$  邻域代表模式、  $LBP_{p,r}^{n_u}$  邻域主梯度代表模式的特征维数 3 960 维,CS\_RBP 单点、CS\_RBP 邻域代表模式、CS\_RBP 梯度邻域主梯度代表模式的特征维数均为 10 608 维.HOG 特征的 block 单元由 8×8 像素构成,2×2 个 block 构成一个 cell 单元,最后提取的特征维数为 20 736 维.表 3~表 5 中给出了不同方法在 3 个数据集上的测试结果,其中,占比是指即训练样本集占总数据集图片的比例.

由表 3~表 5 可以看出:在 3 个数据集上,所有方法被分成 5 组,其中,第 1 组为表中第 2 行,第 2 组为表中第 3 行、第 4 行,第 3 组为表中第 5 行~第 8 行,第 4 组为表中第 9 行~第 12 行、第 5 组为表中第 13 行~第 16 行.在传统 LBP 方法组中,  $LBP_{8,2}^{n_u}$  方法的识别率高于  $LBP_{8,1}^{n_u}$  方法,说明适当增大邻域半径可以捕获更多的邻域信息,对于提高识别率有帮助.在不采用邻域选择策略的单点模式对方法组中,在 3 个数据集上,  $LBP_{p,r}^{n_u}$  单点与 CS\_RBP 单点的方法普遍表现较差.原因在于:一是图像 ROI 区域中心位置估算存在误差,二是图像本身噪声的干扰.此时,直接将该点的编码值作为模式值难以获得稳定的特征表达.但是 CS\_RBP 单点方法性能仍高于  $LBP_{8,2}^{n_u}$  单点方法,原因在于 CS\_RBP 可以编码 16 种模式,多于  $LBP_{p,r}^{n_u}$  方法的 10 种模式,能表达图像更多的细节信息.此外,在 3 个数据集上,CS\_RBP 的邻域代表模式方法与邻域主梯度代表模式方法的识别率也均高于  $LBP_{p,r}^{n_u}$  的邻域代表模式方法与邻域主梯度代表模式方法,充分说明 CS\_RBP 方法特征的鉴别能力强于传统 LBP 特征模式.表明采用局部径向坐标变换技术后,同样是基于邻域选择策略的主模式,具有 16 个模式种类的 CS\_RBP 优于只有 10 个模式种类  $LBP_{p,r}^{n_u}$ .同时,在  $LBP_{p,r}^{n_u}$ 、CS\_RBP 方法中,基于邻域的代表模式均远高于其对应的单点类的方法,表明基于邻域选择策略的代表模式鲁棒性强,能够很好地克服图像 ROI 区域中心位置估算误差以及图像本身的噪声与干扰的影响.

Table 3 Euro coin recognition accuracy

表 3 欧元硬币数据集识别率

方法	占比			
		2%	4%	8%
HOG		57.46% (4844/8430)	67.73% (5710/8430)	83.44% (7034/8430)
$LBP_{8,1}^{n_u}$		94.39% (7957/8430)	99.42% (8381/8430)	99.85% (8417/8430)
$LBP_{8,2}^{n_u}$		94.71% (7984/8430)	99.62% (8398/8430)	99.99% (8429/8430)
$LBP_{8,1}^{n_u}$ 单点		48.48% (4087/8430)	49.50% (4173/8430)	51.61% (4351/8430)
$LBP_{8,2}^{n_u}$ 单点		48.48% (4087/8430)	49.50% (4173/8430)	51.65% (4354/8430)
CS_RBP <sub>4,1</sub> 单点		48.49% (4088/8430)	49.51% (4174/8430)	51.60% (4350/8430)
CS_RBP <sub>4,2</sub> 单点		48.49% (4088/8430)	49.51% (4174/8430)	66.75% (5627/8430)
$LBP_{8,1}^{n_u}$ 邻域代表模式		91.03% (7674/8430)	96.47% (8132/8430)	97.86% (8250/8430)
$LBP_{8,2}^{n_u}$ 邻域代表模式		90.64% (7641/8430)	95.37% (8040/8430)	97.94% (8256/8430)
CS_RBP <sub>4,1</sub> 邻域代表模式		95.18% (8024/8430)	99.02% (8347/8430)	100% (8430/8430)
CS_RBP <sub>4,2</sub> 邻域代表模式		95.79% (8075/8430)	98.71% (8321/8430)	99.70% (8405/8430)
$LBP_{8,1}^{n_u}$ 邻域主梯度代表模式		83.04% (7000/8430)	89.94% (7582/8430)	94.16% (7938/8430)
$LBP_{8,2}^{n_u}$ 邻域主梯度代表模式		84.38% (7113/8430)	94.85% (7996/8430)	98.40% (8295/8430)
CS_RBP <sub>4,1</sub> 邻域主梯度代表模式		87.13% (7345/8430)	97.90% (8253/8430)	99.82% (8415/8430)
CS_RBP <sub>4,2</sub> 邻域主梯度代表模式		72.25% (6091/8430)	89.91% (7579/8430)	97.78% (8243/8430)

Table 4 QQ expression recognition accuracy

表 4 QQ 表情数据集识别率

方法	占比	6%	11%	16%
HOG		34.59% (1295/3744)	63.2746% (2369/3744)	89.98% (3369/3744)
$LBP_{8,1}^{\mu_2}$		50.13% (1877/3744)	85.87% (3215/3744)	94.52% (3539/3744)
$LBP_{8,2}^{\mu_2}$		62.23% (2330/3744)	94.98% (3556/3744)	97.97% (3668/3744)
$LBP_{8,1}^{\mu_2}$ 单点		15.20% (569/3744)	26.68% (999/3744)	40.65% (1522/3744)
$LBP_{8,2}^{\mu_2}$ 单点		16.19% (606/3744)	28.04% (1050/3744)	45.01% (1685/3744)
CS_RBP <sub>4,1</sub> 单点		15.92% (596/3744)	30.48% (1141/3744)	49.44% (1851/3744)
CS_RBP <sub>4,2</sub> 单点		15.22% (570/3744)	29.59% (1108/3744)	48.90% (1831/3744)
$LBP_{8,1}^{\mu_2}$ 邻域代表模式		33.31% (1247/3744)	54.06% (2024/3744)	69.31% (2595/3744)
$LBP_{8,2}^{\mu_2}$ 邻域代表模式		37.77% (1414/3744)	68.64% (2570/3744)	83.23% (3116/3744)
CS_RBP <sub>4,1</sub> 邻域代表模式		92.65% (3469/3744)	98.72% (3696/3744)	100% (3744/3744)
CS_RBP <sub>4,2</sub> 邻域代表模式		91.85% (3439/3744)	98.21% (3677/3744)	100% (3744/3744)
$LBP_{8,1}^{\mu_2}$ 邻域主梯度代表模式		26.82% (1004/3744)	55.74% (2087/3744)	77.38% (2897/3744)
$LBP_{8,2}^{\mu_2}$ 邻域主梯度代表模式		25% (936/3744)	58.07% (2174/3744)	81.94% (3068/3744)
CS_RBP <sub>4,1</sub> 邻域主梯度代表模式		93.94% (3517/3744)	98.10% (3673/3744)	100% (3744/3744)
CS_RBP <sub>4,2</sub> 邻域主梯度代表模式		93.03% (3483/3744)	97.77% (3660/3744)	100% (3744/3744)

Table 5 Car logo recognition accuracy

表 5 车标数据集识别率

方法	占比	15%	30%	45%
HOG		100% (1500/1500)	100% (1500/1500)	100% (1500/1500)
$LBP_{8,1}^{\mu_2}$		80.67% (1210/1500)	91.87% (1378/1500)	94.93% (1424/1500)
$LBP_{8,2}^{\mu_2}$		84.93% (1274/1500)	94.47% (1417/1500)	96.6% (1449/1500)
$LBP_{8,1}^{\mu_2}$ 单点		62.27% (934/1500)	48.8% (732/1500)	61.93% (929/1500)
$LBP_{8,2}^{\mu_2}$ 单点		39.8% (597/1500)	57.13% (857/1500)	68.4% (1026/1500)
CS_RBP <sub>4,1</sub> 单点		66.33% (995/1500)	78.00% (1170/1500)	84.87% (1273/1500)
CS_RBP <sub>4,2</sub> 单点		70.20% (1053/1500)	80.60% (1209/1500)	85.93% (1289/1500)
$LBP_{8,1}^{\mu_2}$ 邻域代表模式		37.13% (557/1500)	51.53% (773/1500)	63.13% (947/1500)
$LBP_{8,2}^{\mu_2}$ 邻域代表模式		55.4% (831/1500)	68.4% (1026/1500)	77.73% (1166/1500)
CS_RBP <sub>4,1</sub> 邻域代表模式		83.87% (1258/1500)	91.27% (1369/1500)	94.07% (1411/1500)
CS_RBP <sub>4,2</sub> 邻域代表模式		83.93% (1259/1500)	92.53% (1388/1500)	94.93% (1424/1500)
$LBP_{8,1}^{\mu_2}$ 邻域主梯度代表模式		62.27% (934/1500)	75.4% (1131/1500)	88.33% (1325/1500)
$LBP_{8,2}^{\mu_2}$ 邻域主梯度代表模式		71.33% (1070/1500)	81.6% (1224/1500)	83.27% (1249/1500)
CS_RBP <sub>4,1</sub> 邻域主梯度代表模式		88.93% (1334/1500)	94.13% (1412/1500)	96.53% (1448/1500)
CS_RBP <sub>4,2</sub> 邻域主梯度代表模式		89.87% (1348/1500)	95.07% (1426/1500)	97.07% (1456/1500)

LBP 方法比较适合提取纹理特征,但是在纹理信息比较缺乏的 QQ 表情和车标数据集中,对应的 LBP 方法的性能远低于本文的方法.这也说明本文方法适应性强,对于纹理丰富和纹理缺乏的图像均能适用.值得注意的是:邻域半径增大后,很多时候 CS\_RBP<sub>4,2</sub> 并没有取得 CS\_RBP<sub>4,1</sub> 更好的性能,反而还有所下降.原因是局部径向坐标变换的旋转不变二值模式的近似计算在半径增大后,计算误差会增加,尤其是当图像邻域中含有细节纹理信息时,例如在含有较多纹理信息的硬币数据集和因为色差产生的具有纹理变化的 QQ 表情数据集测试时.

由表 3~表 5 可知:在训练集占比为 8% 的欧元硬币数据集与占比为 16% 的 QQ 表情数据集上,能够达到的最好识别率 100% 均为本文方法所取得;只有在车标数据集中,获得的最好识别率 100% 由 HOG 方法取得.同时,我们可观察到:只需要较少比例的训练样本,例如欧元硬币数据集中训练集占比为 8%、QQ 表情数据集中训练集

占比 16%,本文方法均达到 100%的识别率.但是对于车标数据集,尽管车标图案不旋转,但是由于图像在自然场景下获取,背景、照明和视角均有较大变化,因此仍然需要较多的训练样本,本文方法才能保证其识别性能.

由于 HOG 特征本身没有考虑旋转不变性问题,因此其对于旋转图案识别效果不理想,特别是在样本数较少的情况下,由于样本很难涵盖各种旋转角度,此时 HOG 方法识别性能较差.例如在硬币数据集中,在训练样本数为 2%时,HOG 特征识别率只有 57.46%,而 LBP 和 CS\_RBP 邻域代表模式能达到 90%以上的识别率;在训练样本数为 4%时,HOG 特征识别率只有 67.73%,而 LBP 和 CS\_RBP 邻域代表模式多数情况能达到 95%以上的识别率;在训练样本数为 8%时,HOG 特征识别率虽然可以达到 83.44%,但是仍远低于 LBP,CS\_RBP 邻域代表模式和邻域主梯度代表模式的识别率.在 QQ 表情图像集中,相似的情况也能在 HOG 方法和 CS\_RBP 邻域主梯度代表模式中观察到.但是对于车标数据集,由于车标图像本身无旋转变化且边缘结构信息显著,此时 HOG 特征表现优异,在训练样本占比 15%,30%,45%这 3 种情况下都得到了 100%的识别准确率.这是因为 HOG 特征描述利用的是梯度信息,其在对抗照明变化上要优于 LBP 和 CS\_RBP 特征方法,但是一旦图像发生旋转,则 HOG 特征的性能难以得到保证,这在欧元硬币的检测和 QQ 表情数据集中已经得到验证.

### 2.2.3 计算效率比较

CS\_RBP 作为二值描述算子,计算原理简单,与其他二值描述算子一样具有高效的优点.表 6 给出了几种代表性的浮点数值型描述子 HOG,SIFT 和 SURF,以及二值描述算子 BRIEF,FREAK 和本文 CS\_RBP 方法在进行单点特征提取所花时间的比较.其中:HOG 单个 Cell 特征提取约耗时 1.9ms;SIFT 和 SURF 算子需要 2.5ms 和 1.4ms;而 CS\_RBP 方法只需 0.045ms,效率是 HOG 的 42 倍、SIFT 的 56 倍、SURF 的 31 倍.由于 CS\_RBP 在寻找对称模式时计算原理比 BRISK 和 FREAK 复杂一些,因此时间消耗上略有增加.CS\_RBP 算法运行环境:CPU 为 i5-4590,4GB 内存,Matlab 2014a.

**Table 6** Comparison single point feature extraction time of different descriptors

表 6 描述子单个点特征提取时间比较

方法	HOG	SIFT	SURF	BRISK	FREAK	CS_RBP
时间(ms)	1.9	2.5	1.4	0.031	0.018	0.045

### 2.2.4 抗旋转性能测试

为验证本文算法特征的抗旋转特性,以德国 2 欧元硬币背面图像为例,如图 12 所示,分别提取初始样本和旋转不同角度后该硬币图像的特征,特征向量之间的相关度由公式(6)计算:

$$P(H_Y, H_M) = \frac{\langle H_Y, H_M \rangle}{\sqrt{\langle H_Y, H_Y \rangle} \sqrt{\langle H_M, H_M \rangle}} \quad (6)$$

其中, $H_Y$  和  $H_M$  分别表示初始样本图片和旋转后图片的特征向量, $\langle \cdot, \cdot \rangle$  表示向量内积.

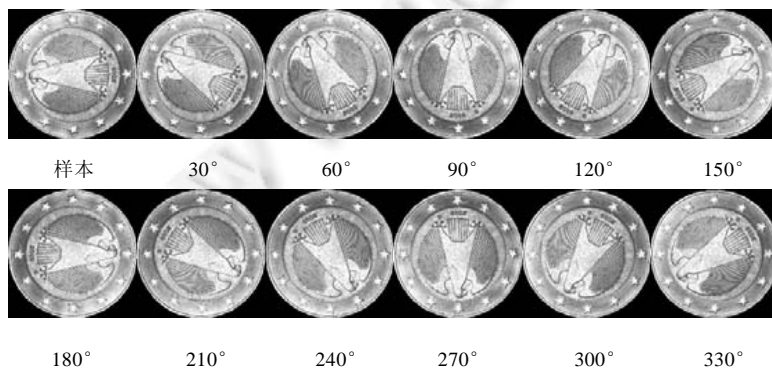


Fig.12 Test picture rotation schematic

图 12 测试图片旋转示意图

$P(H_Y, H_M)$ 表示  $H_Y$  和  $H_M$  的相关系数,可以衡量特征的抗旋转能力,平均相关度越大且标准差越小,则抗旋转性越强.每次旋转 15 度,我们计算了原始图像和旋转后图像的特征相关度,表 7 给出了本文 4 种方法和 LBP,HOG 特征在图像旋转后的相关度值.

**Table 7** Average correlation of the features after rotation of the sample images

表 7 样本图片旋转后的特征平均相关度

旋转角度	CS_RBP <sub>4,1</sub> 邻域代表模式	CS_RBP <sub>4,2</sub> 邻域代表模式	CS_RBP <sub>4,1</sub> 邻域 主梯度代表模式	CS_RBP <sub>4,2</sub> 邻域 主梯度代表模式	HOG	LBP <sub>8,1</sub> <sup>riu2</sup>	LBP <sub>8,2</sub> <sup>riu2</sup>
15°	0.82	0.89	0.90	0.79	0.79	1	1
30°	0.82	0.91	0.87	0.77	0.72	1	1
45°	0.81	0.90	0.89	0.78	0.66	1	1
60°	0.82	0.87	0.85	0.75	0.61	1	1
75°	0.83	0.92	0.86	0.76	0.60	1	1
90°	0.98	0.98	0.98	0.95	0.59	1	1
105°	0.83	0.88	0.89	0.79	0.58	1	1
120°	0.82	0.90	0.87	0.77	0.55	1	1
135°	0.82	0.89	0.88	0.77	0.55	1	1
150°	0.82	0.87	0.85	0.74	0.57	1	1
165°	0.83	0.91	0.85	0.77	0.56	1	1
180°	0.97	0.97	0.97	0.93	0.58	1	1
195°	0.83	0.88	0.90	0.78	0.64	1	1
210°	0.82	0.89	0.87	0.77	0.67	1	1
225°	0.83	0.89	0.88	0.76	0.65	1	1
240°	0.83	0.86	0.85	0.74	0.62	1	1
255°	0.83	0.92	0.86	0.75	0.60	1	1
270°	0.99	0.99	0.98	0.95	0.56	1	1
285°	0.83	0.88	0.90	0.79	0.53	1	1
300°	0.82	0.90	0.87	0.79	0.53	1	1
315°	0.83	0.91	0.89	0.76	0.56	1	1
330°	0.82	0.87	0.85	0.73	0.63	1	1
345°	0.83	0.93	0.85	0.76	0.77	1	1
平均值	0.845	0.905	0.885	0.789	0.614	0	0
标准差	0.054	0.035	0.040	0.063	0.071	0	0

从表 7 可知,本文的 4 种算法在每个旋转角度的相关度均比 HOG 特征高,同时标准差也较小.本文 4 种算法的平均相关度分别为 0.845,0.905,0.885,0.789,远高于 HOG 特征的 0.614,其中,CS\_RBP<sub>4,2</sub> 邻域代表模式表现最为稳定,平均相关度值达到 0.905,标准差 0.035.LBP 特征模式我们采用的是旋转不变模式,因此其本身具有旋转不变性,在所有旋转角度情况下,其特征相关度均为 1.尽管如此,但是 LBP 旋转不变模式信息损失较大,不利于图像细节的描述,因此特征的鉴别能力不足.而本文方法在尽可能保留抗旋转描述特性的同时,利用空间对称位置结构信息有效提高了特征的鉴别能力,从而达到抗旋转高鉴别二值模式特征提取的目的.

2.2.5 原始特征计算方法和近似计算方法对比

在基于局部径向坐标变换的旋转不变二值模式计算时,我们采用了近似计算方法.为了验证近似方法的可靠性,与原始算法进行了对比测试.考虑到原始特征计算方法相当耗时,我们只选取了部分欧元硬币进行实验,其中,2 欧元硬币 580 张,1 欧元硬币 552 张,50 欧分硬币 252 张,20 欧分硬币 312 张,10 欧分硬币 312 张,实验结果见表 8.

**Table 8** Initial feature calculation method is compared with the recognition result of approximate calculation method in euro coin recognition

表 8 原始特征计算方法与近似计算方法在欧元硬币识别上的识别结果对比

原始 CS_RBP <sub>4,1</sub> 邻域 代表模式	近似 CS_RBP <sub>4,1</sub> 邻域 代表模式	原始 CS_RBP <sub>4,2</sub> 邻域 代表模式	近似 CS_RBP <sub>4,2</sub> 邻域 代表模式	原始 CS_RBP <sub>4,2</sub> 邻域主梯度 代表模式	近似 CS_RBP <sub>4,2</sub> 邻域主梯度 代表模式	原始 CS_RBP <sub>4,2</sub> 邻域主梯度 代表模式	近似 CS_RBP <sub>4,2</sub> 邻域主梯度 代表模式
100% (2008/2008)	100% (2008/2008)	99.65% (2001/2008)	100% (2008/2008)	99.90% (2006/2008)	99.85% (2005/2008)	100% (2008/2008)	98.06% (1969/2008)

从表8可看出,4种情况下的原始特征计算方法与近似计算方法上在最终的分类精度上相差无几;而近似计算方法由于避免了大量的浮点数插值计算,从而大大提高了计算效率;原始特征计算方法对  $200 \times 200$  大小的图像进行的特征提取耗时约 30s,而近似计算办法耗时仅需 1s.

### 3 总结

针对圆形图案中心对称的特点,本文提出一种基于空间对称结构模式对的图像二值特征描述方法.由于引入了局部径向变换技术,在抗旋转特性的同时,使得能够表达的基本模式达到 16 种,多于  $LBP_{p,r}^{riu_2}$  的 10 种模式,因此能够更多地描述图像的微观结构细节信息;并且通过组成对称模式对,充分利用了圆形图像的空间结构信息,因此提升了特征的鉴别能力.在具有旋转情况的欧元硬币和 QQ 表情数据集上,均取得了优于原 LBP 特征和 HOG 特征提取方法的识别性能.由于算法专门针对旋转的圆形物体图像识别问题,因此 CS\_RBP 方法在抗旋转能力上具有很强的鲁棒性.在训练集占比为 8% 的欧元硬币数据集与训练集占比 16% QQ 的表情数据集中,本文方法均达到 100% 的识别率.说明本文方法只需少量的代表性训练样本即可达到良好的性能,不需要存储大量的不同旋转角度的训练样本集.此外,CS\_RBP 作为一种高效的二值描述算子,具有和其他二值描述算子一样的优点.单点计算时间只需 0.045ms,效率是 HOG 的 42 倍、SIFT 算子的 56 倍、SURF 算子的 31 倍.

**作者注** 本文是我们于 2017 年 5 月 24 日投到《软件学报》的论文,该文是湘潭大学张东波老师指导的 2018 届(2018 年 6 月)毕业的研究生陈红磊(本文第二作者)的硕士论文《基于计算机视觉的欧元硬币识别》工作成果的一部分.特此说明.

### References:

- [1] Lowe DG. Distinctive image features from scale-invariant keypoints. *Int'l Journal of Computer Vision*, 2004,60(2):91–110. [doi: 10.1023/B:VISI.0000029664.99615.94]
- [2] Bay H, Tuytelaars T, Van Gool L. SURF: Speeded up robust features. In: *Proc. of the European Conf. on Computer Vision*. 2006. 404–417. [doi: 10.1007/11744023\_32]
- [3] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005,27(10):1615–1630. [doi: 10.1109/TPAMI.2005.188]
- [4] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2005. 886–893. [doi: 10.1109/CVPR.2005.177]
- [5] Song KC, Yan YH, Chen WG, Zhang X. Research and perspective on local binary pattern. *Acta Automatica Sinica*, 2013,39(6): 730–744.
- [6] Calonder M, Lepetit V, Ozuysal M, Trzcinski T, Strecha C, Fua P. BRIEF: Computing a local binary descriptor very fast. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012,34(7):1281–1298. [doi: 10.1109/TPAMI.2011.222]
- [7] Leutenegger S, Chli M, Siegwart RY. BRISK: Binary robust invariant scalable keypoints. In: *Proc. of the Int'l Conf. on Computer Vision*. 2011. 2548–2555. [doi: 10.1109/ICCV.2011.6126542]
- [8] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In: *Proc. of the Int'l Conf. on Computer Vision*. 2011. 2564–2571. [doi: 10.1109/ICCV.2011.6126544]
- [9] Alahi A, Ortiz R, Vandergheynst P. FREAK: Fast retina keypoint. In: *Proc. of the Computer Vision and Pattern Recognition*. 2012. 510–517. [doi: 10.1109/CVPR.2012.6247715]
- [10] Fan B, Wu F, Hu Z. Rotationally invariant descriptors using intensity order pooling. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012,34(10):2031–2045. [doi: 10.1109/TPAMI.2011.277]
- [11] Guan W, You S. Robust image matching with line context. In: *Proc. of the British Machine Vision Conf.* 2013. [doi: 10.5244/C.27.34]

- [12] Takacs G, Chandrasekhar V, Tsai SS, Chen DM, Grzeszczuk R, Girod B. Unified real-time tracking and recognition with rotation-invariant fast features. In: Proc. of the Computer Vision and Pattern Recognition. 2010. 934–941. [doi: 10.1109/CVPR.2010.5540116]
- [13] Chandrasekhar V, Takacs G, Chen DM, Tsai SS, Grzeszczuk R, Girod B. CHoG: Compressed histogram of gradients a low bit-rate feature descriptor. In: Proc. of the Computer Vision and Pattern Recognition. 2009. 2504–2511. [doi: 10.1109/CVPR.2009.5206733]
- [14] Wang Z, Fan B, Wu F. Local intensity order pattern for feature description. In: Proc. of the Int'l Conf. on Computer Vision. 2011. 603–610. [doi: 10.1109/ICCV.2011.6126294]



张东波(1973—),男,湖南邵阳人,博士,教授,CCF 专业会员,主要研究领域为模式识别,图像处理,机器学习.



汤红忠(1979—),女,博士,副教授,主要研究领域为图像处理与模式识别,字典学习及其稀疏表示.



陈红磊(1991—),男,硕士生,主要研究领域为模式识别,图像处理,机器学习.



许海霞(1978—),女,博士,副教授,主要研究领域为模式识别,机器视觉.



文登伟(1990—),男,硕士生,主要研究领域为模式识别,图像处理,机器学习.

www.jos.org.cn