

# 基于改进二值模式的图书文档图像分类

张 敏

(河南理工大学 测绘与国土信息工程学院, 河南 焦作 454003)

**摘要:** 在分析针对局部二值模式进行降维方法的基础上, 提出了一种改进的局部二值模式描述符, 并用于图书文档图像分类。新方法首先将原局部邻域划分为多个4-正交邻域, 然后统计4-正交邻域二值化后所包含的“1”的个数作为特征, 最后融合所有4-正交邻域特征进行图像分类。采用广泛应用的纹理图像库、前视红外目标图像库和图书文档图像库进行实验, 结果表明, 新方法的特征维数不但明显降低, 而且还取得了较高的分类准确率。

**关键词:** 图像分类; 局部二值模式; 纹理分析; 降维

**中图分类号:** TN219      **文献标识码:** A      **文章编号:** 1001-8891(2014)10-0827-05

## Document Image Classification Based on Improved Local Binary Patterns

ZHANG Min

(School of Surveying & Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China)

**Abstract:** Based on the analysis of the methods to reduce the dimensions of the local binary pattern(LBP), a new operator called the orthogonal combination of local binary number (denoted as OC-LBN) is proposed for document image classification. Firstly, the local neighborhood is divided into different 4-orthogonal-neighbors, and the binary number of “1” in each 4-orthogonal-neighbor is used as its feature. Then, the features of all the 4-orthogonal-neighbor are combined together as region description. Experimental results obtained from texture, forward-Looking infrared and document image databases demonstrate that the proposed method can get the best performance of the methods mentioned in the paper.

**Key words:** image classification, local binary pattern, texture analysis, dimensionality reduction

## 0 引言

随着数字图书馆技术的发展, 数字化信息呈现几何级数的增长趋势, 其中尤其以数字化图像、数字化视频、数字化音频数据的增长为主导。因此, 如何快速、高效地从数字图书馆中找出用户所需的资源就成为现代图书馆研究的热点和关键技术。近年来, 针对图像资源的检索, 基于内容的图像检索技术(Content Based on Image Retrieval, CBIR)得到了广泛研究<sup>[1]</sup>。另外, 如果事先对所检索图像资源进行有效分类, 检索时按照分类进行将大大缩减检索所需时间。在图像特征的描述中, 颜色、形状及纹理是最常用的特征, 本文基于纹理特征, 研究有效的图像分类方法。

作为简单有效的纹理分析方法, 近年来, 局部二值模式(Local Binary Pattern, LBP)<sup>[2]</sup>得到了广泛研

究。由于LBP原理及计算简单, 并融合了纹理的结构与统计特征, 因此在图像处理与模式识别领域得到了广泛应用<sup>[3-5]</sup>。但原始LBP维数较高(对于8-邻域, 维数达256维), 为了有效降低维数, 扩展LBP的应用范围, 各种针对LBP的降维方法也被提了出来。Ojala等<sup>[6]</sup>将LBP扩展至适合任意的圆形邻域, 同时提出了一致局部二值模式U-LBP(Uniform LBP)来降低维数(对于8-邻域, 将LBP由256维降为59维); Heikkilä等<sup>[7]</sup>提出了中心对称局部二值模式(Center-Symmetric LBP, CS-LBP)来降低维数(对于8-邻域, 将LBP由256维降为16维); 在CS-LBP的基础上, 毋小省等提出了方向局部二值模式D-LBP(Direction LBP)<sup>[8]</sup>及局部边缘二值模式(Local Edge Binary Pattern, LEBP)<sup>[9]</sup>, 在实现降维的同时, 进一步提高了CS-LBP的分辨能力; Zhu等<sup>[10]</sup>提出了

收稿日期: 2014-03-03; 修订日期: 2014-08-22.

作者简介: 张敏(1962-), 女, 河南南阳人, 馆员, 主要研究方向是图书文档识别、图像检索技术等。E-mail: zhangmin@hpu.edu.cn.

基金项目: 河南省国际合作项目, 编号: 134300510057.

OC-LBP(Orthogonal Combination of LBP)来降低 LBP 的维数(对于 8-邻域, 将 LBP 由 256 维降为 32 维)。

为进一步降低 LBP 的特征维数、同时提高 LBP 描述符的性能, 在 OC-LBP 的基础上, 本文提出了一种改进的局部二值模式 OC-LBN (Orthogonal Combination of Local Binary Number)。新方法采用与 OC-LBP 完全不同的编码方式, 通过统计局部邻域二值化后值为“1”像素的数目作为局部邻域特征。

文中首先采用纹理图像与前视红外目标图像进行实验, 以验证新算法的性能; 在此基础上, 将新描述符用于图书文档图像分类。结果表明, 本文 OC-LBN 在保证计算简单的前提下, 其性能明显优于 U-LBP<sup>[6]</sup>、CS-LBP<sup>[7]</sup>及 OC-LBP<sup>[10]</sup>3 种描述符, 同时 OC-LBN 的特征维数还低于其它 3 种描述符的特征维数。

## 1 局部二值模式

### 1.1 LBP

局部二值模式 LBP<sup>[2]</sup>基本原理是对于图像 8-邻域, 首先比较中心像素与其邻域像素的灰度值大小, 如果邻域像素灰度值不小于中心像素灰度值, 则该邻域像素赋为 1, 否则赋为 0; 然后将该邻域的二值化结果与一权值模板对应相乘求和, 即得到 LBP 值。

鉴于 LBP 简单有效的特点, Ojala 等<sup>[6]</sup>进一步将其扩展至任意圆形邻域(P, R), 其定义如下:

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(p_i - p_c) \times 2^i \quad (1)$$

式中:  $s(p_i - p_c) = \begin{cases} 1, & \text{if } p_i - p_c \geq 0 \\ 0, & \text{otherwise} \end{cases}$ ;  $p_c$  及  $p_i$  分别为

中心像素和邻域像素的灰度值;  $P$  为邻域像素个数;  $R$  为邻域半径。图 1 给出了 3 种不同类型的邻域, 图 2 给出了 8-邻域的 LBP 计算示例。

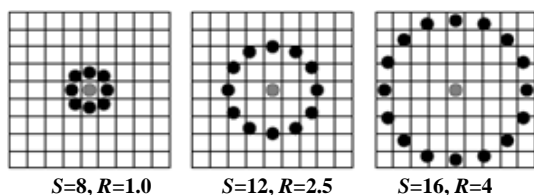


图 1 圆形邻域示例 Fig.1 Examples of circle region

由于 LBP 特征维数高, 达到  $2^P$  维。为此, Ojala 等<sup>[6]</sup>又提出了一致局部二值模式 U-LBP。U-LBP 定义为: 如果局部二值模式中由“0”~“1”或由“1”~“0”的转换次数  $U$  小于或等于 2, 则定义该模式为一一致局部二值模式, 转换次数大于 2 的所有二值模式统一作为一种二值模式。 $U$  的定义如下:

$$U(LBP_{P,R}) = |s(p_{P-1} - p_c) - s(p_0 - p_c)| + \sum_{i=1}^{P-1} |s(p_i - p_c) - s(p_{i-1} - p_c)| \quad (2)$$

如对于二值模式 11111111, 转换次数  $U=0$ ; 对于二值模式 11000001, 转换次数  $U=2$ ; 对于 10101011, 转换次数  $U=6$ 。基于该定义, 一致局部二值模式 U-LBP 将 LBP 的维数从  $2^P$  维降为  $P \times (P-1) + 3$  维。

### 1.2 CS-LBP

为降低 LBP 特征维数, Heikkilä 等<sup>[7]</sup>提出了 CS-LBP 描述符。不同于 LBP 的定义方式, CS-LBP 仅通过与中心像素对称的 4 对像素间的灰度关系来定义。表示为:

$$CS\_LBP_{P,R} = \sum_{i=0}^{P/2-1} s(p_i - p_{i+P/2} - \tau) \times 2^i \quad (3)$$

式中:  $s(p_i - p_{i+P/2} - \tau) = \begin{cases} 1, & \text{if } p_i - p_{i+P/2} \geq \tau \\ 0, & \text{otherwise} \end{cases}$ ;  $\tau$  为预

先设定阈值, 用于判别局部区域的平坦性。显然, CS-LBP 的特征维数仅为  $2^{P/2}$  维。图 2 也给出了 8-邻域的 CS-LBP 计算示例。

### 1.3 OC-LBP

针对 LBP 维数高的问题, Zhu 等<sup>[10]</sup>提出了 OC-LBP 描述符。OC-LBP 首先通过将原局部邻域划分为多个 4-正交邻域, 然后采用与 LBP 相同的方法计算每个 4-正交邻域的局部二值模式值, 最后融合所有 4-正交邻域的二值模式值作为最终的描述。设  $M=P/4$  表示局部邻域被划分为 4-正交邻域的个数,  $m=1, 2, \dots, M$ , OC-LBP 定义如下:

$$\begin{aligned} OC\_LBP1 &= \sum_{i=0}^3 s(p_{M \times i} - p_c) \times 2^i \\ OC\_LBP2 &= \sum_{i=0}^3 s(p_{M \times i+1} - p_c) \times 2^i \\ &\vdots \\ OC\_LBPm &= \sum_{i=0}^3 s(p_{m+M \times i-1} - p_c) \times 2^i \\ &\vdots \\ OC\_LBPm &= \sum_{i=0}^3 s(p_{M+M \times i-1} - p_c) \times 2^i \end{aligned} \quad (4)$$

$$OC\_LBP = [OC\_LBP1, \dots, OC\_LBPm, \dots, OC\_LBPm] \quad (5)$$

显然 OC-LBP 的特征维数为  $P \times 4$ , 图 2 也给出了针对 8-邻域的 OC-LBP 计算示例。

## 2 改进二值模式

对于 U-LBP<sup>[6]</sup>及 CS-LBP<sup>[7]</sup>来说, 虽然这两种方法可以有效降低 LBP 描述符的特征维数, 但 U-LBP 及 CS-LBP 的特征的维数仍随着邻域尺寸的升高而迅速升高; D-LBP<sup>[8]</sup>及 LEBP<sup>[9]</sup>两种特征具有同 CS-LBP 完全一样的维数, 并且这两种特征的维数对邻域尺寸的变化也很敏感。对于 OC-LBP<sup>[10]</sup>来说, 该方法通过 4-正交邻域的划分方法, 既有效降低了原 LBP 描述符的

特征维数,也降低了邻域尺寸变化对特征维数的影响。

为此,本文借助于 OC-LBP 的 4-正交邻域划分方法,提出了改进的局部二值模式 OC-LBN。新方法在对二值化后的局部邻域进行编码时,采用了与 LBP、CS-LBP 及 OC-LBP 完全不同的方式,通过统计局部邻域二值化后值为“1”像素的数目作为局部邻域特征。新方法在进一步降低特征维数的同时,也进一步降低了邻域尺寸变化对特征维数的影响程度。OC-LBN 实现过程如下:

1) 按照与 OC-LBP<sup>[10]</sup>相同的方法,将局部邻域划分为多个 4-正交邻域,并进行二值化;

2) 在计算每个 4-正交邻域的局部二值模式时,不同于 OC-LBP 的编码方式,OC-LBN 仅仅统计每个二值化后的 4-正交邻域中所包含“1”的个数作为二值模式的值;

3) 融合多个 4-正交邻域的 OC-LBN 作为最终特征。

设  $M=P/4$  表示局部邻域被划分为 4-正交邻域的个数,  $m=1, 2, \dots, M$ , OC-LBN 定义如下:

$$\begin{aligned} \text{OC\_LBN1} &= \sum_{i=0}^3 s(p_{Mxi} - p_c) \\ \text{OC\_LBN2} &= \sum_{i=0}^3 s(p_{Mxi+1} - p_c) \\ &\vdots \\ \text{OC\_LBNm} &= \sum_{i=0}^3 s(p_{m+Mxi-1} - p_c) \\ &\vdots \\ \text{OC\_LBNM} &= \sum_{i=0}^3 s(p_{M+Mxi-1} - p_c) \end{aligned} \quad (6)$$

$$\text{OC\_LBN} = [\text{OC\_LBN1}, \dots, \text{OC\_LBNm}, \dots, \text{OC\_LBNM}] \quad (7)$$

显然 OC-LBN 的特征维数为  $M=P/4 \times 5$  维,图 2 也给出了针对 8-邻域的 OC-LBN 计算示例。

### 3 实验结果

为验证 OC-LBN 的性能,我们首先进行了纹理图像和前视红外目标图像的分类实验,然后针对图书文档图像进行分类实验,并将其同 U-LBP<sup>[6]</sup>、CS-LBP<sup>[7]</sup> 及 OC-LBP<sup>[10]</sup> 3 种方法进行了比较。

为了充分验证不同方法的性能,在对邻域( $P, R$ )的选取上,实验中分别选择(4, 1)、(8, 1)、(16, 2)及(24, 3)4 种类型的邻域;对于 CS-LBP, 实验中取  $\tau=5$ 。

在进行图像分类中,采用 K-NN 方法(实验中取  $K=1$ )实现纹理图像分类;采用  $\chi^2$  距离计算特征间的相似性。设  $H$  及  $B$  表示两幅图像的二值模式特征直方图,  $\chi^2$  距离表示如下:

$$d(H, B) = \sum_{k=1}^{\text{dim}} \frac{(h_k - b_k)^2}{h_k + b_k} \quad (8)$$

式中: dim 表示特征维数。

#### 3.1 特征维数比较

为了比较 4 种描述的降维效果,表 1 给出了 4 种方法在 4 种不同的邻域条件下特征维数的对比。从表 1 中可以看出:①除(4, 1)邻域外,OC-LBN 的特征维数远小于其它 3 种特征的特征维数;②随着圆形邻域尺寸的增大,CS-LBP 的特征维数受邻域尺寸变化的影响最大,其次是 U-LBP, 然后是 OC-LBP, OC-LBN 的维数变化受邻域尺寸变化的影响最小。通过比较可以看出,新描述符 OC-LBN 不但具有较低的特征维数,并且特征维数的变化受邻域尺寸的影响也最小,因此更适合在高分辨率下作为纹理图像特征。

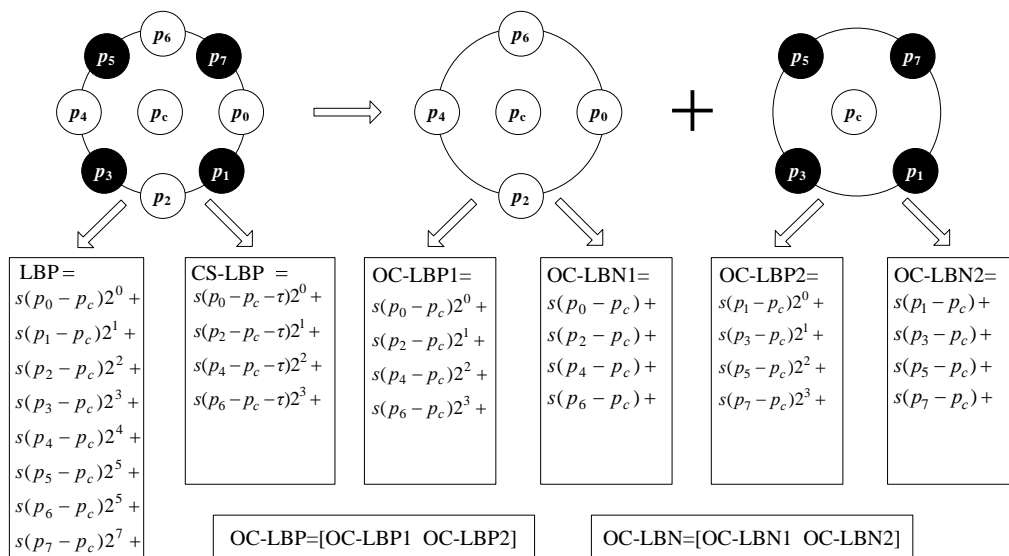


图 2 LBP、CS-LBP、OC-LBP 与 OC-LBN 计算示例 Fig.2 The example of computing LBP, CS-LBP, OC-LBP and OC-LBN

### 3.2 基于 Outex 图像库

对于 Outex 纹理库<sup>[11]</sup>, 我们选择 Outex\_TC\_0010 (TC10) 和 Outex\_TC\_0012 (TC12) 作为实验平台, TC10 和 TC12 均包含 24 类纹理图像, 这些纹理图像是在 3 种不同的光源条件(“horizon”、“inca”及“t184”)下得到, 每类图像包含 9 种不同的旋转角度(0°, 5°, 10°, 15°, 30°, 45°, 60°, 75°, 90°), 每个角度下均有 20 幅 128×128 的图像。对于 TC10 来说, 实验中选择光照条件为 “inca”、旋转角度为 0°的(24×20)幅图像作为训练图像, 其它 8 个旋转角度下光照条件为 “inca” 的 (24×20×8) 幅图像作为测试图像进行分类实验。对于 TC12 图像库来说, 选择光照条件为 “inca” 的所有图像作为训练图像, 其它 “horizon” 及 “t184” 两种光照条件下的所有图像作为测试图像进行分类实验。表 2 给出了 U-LBP、CS-LBP、OC-LBP 与 OC-LBN 4 种方法的分类结果比较, 表中 “t” 代表光照条件 “t184”, “h” 代表光照条件 “horizon”, “-” 表示计算溢出, “avg” 表示在 TC10 与 TC12 及 4 种不同领域条件下的平均分类准确率。

从表 2 中可以明显看出: ①仅仅对于 TC12 纹理库, 在(4, 1)及(24, 3)两种邻域条件下, U-LBP 的分类效果优于 OC-LBN, 但两种条件下 U-LBP 的特征维数分别为 OC-LBN 特征维数的 3 倍和 18.5 倍, 明显高于

OC-LBN 的特征维数; ②在其它所有条件下(包括不同邻域尺寸及 TC10 与 TC12 两个纹理库), OC-LBN 均以最低的特征维数取得最好的分类效果; ③对于所有 4 种邻域及 TC10 和 TC12 两个纹理库来说, OC-LBN 的平均分类准确率 “avg” 明显高于 U-LBP、CS-LBP 及 OC-LBP 3 种方法。

### 3.3 基于前视红外目标图像库

文献[12]对经典红外目标识别的度量方法进行了总结, 文中将局部二值模式特征应用于前视红外目标识别。实验图像库从文献[13]的 FLIR 库中选择 10 类、每类 100 幅共 1000 幅红外目标图像构成, 图像的分辨率为 40×75 像素。实验中, 我们分别在每类中随机选择 10、20、40、60、80 幅图像作为训练图像, 剩余的 90、80、60、40、20 幅图像作为测试图像进行分类实验, 同时将(4, 1)、(8, 1)与(16, 2) 3 种圆形邻域的特征组合在一起作为最终的图像特征。随机进行 10 次实验, U-LBP、CS-LBP、OC-LBP 与 OC-LBN 4 种方法的分类准确率如表 3 所示, 表中 “avg” 表示在 4 种不同领域条件下的平均分类准确率。可以看出, 4 种方法的平均分类准确率差别不大, U-LBP 最高, 其次是 OC-LBP 及 OC-LBN, 但在 3 种圆形邻域特征组合的情况下, U-LBP 特征维数为 OC-LBN 的 9 倍, OC-LBP 特征维数为 OC-LBN 的 3 倍。

表 1 U-LBP、CS-LBP、OC-LBP 与 OC-LBN 特征维数比较

Table 1 The dimension comparison of U-LBP, CS-LBP, OC-LBP and OC-LBN

	特征维数			
	(P,R)=(4,1)	(P,R)=(8,1)	(P,R)=(16,2)	(P,R)=(24,3)
U-LBP	15	59	243	555
CS-LBP	4	16	256	4096
OC-LBP	16	32	64	96
OC-LBN	5	10	20	30

表 2 针对 Outex 图像库的实验结果

Table 2 Experimental results for Outex database

													%	
	(P,R)=(4,1)			(P,R)=(8,1)			(P,R)=(16,2)			(P,R)=(24,3)				
	TC10	TC12		TC10	TC12		TC10	TC12		TC10	TC12			“avg”
		“t”	“h”		“t”	“h”		“t”	“h”		“t”	“h”		
U-LBP	55.05	56.46	56.32	53.93	55.95	58.06	53.20	58.40	59.26	51.85	56.97	56.71	56.01	
CS-LBP	30.47	24.77	25.42	43.46	36.76	41.34	47.34	49.54	51.64	44.74	-	-	39.55	
OC-LBP	53.18	52.73	53.63	53.36	54.86	56.67	52.55	57.18	58.01	45.10	51.97	52.43	53.47	
OC-LBN	67.55	43.52	45.81	73.59	56.11	58.10	69.27	66.00	64.12	57.42	56.83	53.61	59.33	

### 3.4 基于图书文档图像库

该图像库包含 50 类图书封面图像, 每类包含 30 幅具有不同旋转角度以及受噪声感染的图像, 图像的分辨率为  $640 \times 480$ 。实验中, 我们分别选择每类中的前 20 幅图像作为训练图像, 剩余的 10 幅图像作为测试图像进行分类实验, U-LBP、CS-LBP、OC-LBP 与 OC-LBN 4 种方法的分类准确率如表 4 所示, 表中“avg”表示在 4 种不同邻域条件下的平均分类准确率。

从表 4 可以看出: 1) 在(4, 1)邻域下, U-LBP 及 OC-LBP 的分类效果优于 OC-LBN, 但它们的特征维数却达到 OC-LBN 的 3 倍; 2) 在其它邻域下, OC-LBN 均以最低的特征维数取得了最好的分类效果; 3) 对于所有 4 种邻域来说, OC-LBN 的平均分类准确率“avg”明显高于 U-LBP、CS-LBP 及

OC-LBP 3 种方法。

### 4 结束语

在介绍局部二值模式 LBP 及其常用降维方法的基础上, 结合 OC-LBP, 提出了一种新的局部二值模式降维方法 OC-LBN。一方面, OC-LBN 的维数远小于文中所述的 U-LBP、CS-LBP 及 OC-LBP 3 种描述符 (仅对于 4-邻域, OC-LBN 的特征维数比 CS-LBP 特征维数多 1 维); 同时, 随着邻域尺寸的扩大, OC-LBN 的特征维数也不会明显增加, 而其它 3 种方法的维数受邻域尺寸的影响很大。实验中, 首先采用纹理图像和前视红外图像库进行分类实验, 证明了 OC-LBN 的优良分类性能; 在此基础上, 将 OC-LBN 应用与图书文档分类中, 并取得了较好的分类效果。

表 3 针对前视红外目标图像的实验结果 Table 3 Experimental results for forward-Looking infrared image database

	%					
	10	20	40	60	80	“avg”
U-LBP	28.3	33.6	38.1	46.8	52.3	39.8
CS-LBP	26.1	31.3	36.7	48.3	53.2	39.1
OC-LBP	26.4	35.1	37.4	46.2	52.9	39.6
OC-LBN	25.3	35.2	38.4	45.9	52.6	39.5

表 4 针对图书文档图像的实验结果 Table 4 Experimental results for document image database

	%				
	(P,R)=(4,1)	(P,R)=(8,1)	(P,R)=(16,2)	(P,R)=(24,3)	“avg”
U-LBP	43.0	46.4	48.3	48.8	46.6
CS-LBP	21.6	36.0	51.6	-	36.4
OC-LBP	39.6	43.8	47.3	49.2	45.0
OC-LBN	38.7	47.9	52.8	51.8	47.9

### 参考文献:

- [1] 孙君顶, 赵珊. 图像低层特征提取与检索技术[M]. 北京: 电子工业出版社, 2009.
- [2] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on feature distribution[J]. *Pattern Recognition*, 1996, **29**: 51-59.
- [3] Zhong F, Zhang J. Face recognition with enhanced local directional patterns[J]. *Neurocomputing*, 2013, **119**(7): 375-384.
- [4] Subrahmanyam M, Maheshwari R P, Balasubramanian R. Local maximum edge binary patterns: A new descriptor for image retrieval and object tracking[J]. *Signal Processing*, 2012, **92**(6): 1467-1479.
- [5] 宋克臣, 颜云辉, 陈文辉, 等. 局部二值模式方法研究与展望[J]. *自动化学报*, 2013, **39**(6):730-744.
- [6] Ojala T, Pietikäinen M, Mäenpää 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.
- [7] Heikkilä M, Pietikäinen M, Schmid C. Description of interest regions with local binary patterns [J]. *Pattern Recognition*, 2009, **42**(3): 425-436.
- [8] Chao Zhu, Charles-Edmond Bichot, Liming Chen. Image region description using orthogonal combination of local binary patterns enhanced with color information [J]. *Pattern Recognition*, 2013, **46**(7):1949-1963.
- [9] 毋小省. 基于纹理谱特征的图像检索技术研究[D]. 焦作: 河南理工大学, 2010.
- [10] 毋小省, 孙君顶. 基于局部边缘二值模式的图像检索[J]. *光电子·激光*, 2013, **24**(1): 184-189.
- [11] Ojala T, Maenpää T, Pietikainen M, et al. Outex-new framework for empirical evaluation of texture analysis algorithms[C]//*16th International Conference on Pattern Recognition*, 2002, **1**: 701-706.
- [12] 乔立永, 徐立新, 高敏. 红外目标识别图像复杂度度量方法综述[J]. *红外技术*, 2013, **35**(2):88-96.
- [13] Patel V M, Nasrabadi N M, Chellappa R. Sparsity-motivated automatic target recognition[J]. *Applied optics*, 2011, **50**(10): 1425-1433.