

# 结合四元数拉盖尔矩和三维结构的图像哈希

白杰, 赵琰\*

(上海电力大学电子与信息工程学院, 上海 200090)

**摘要:** 目前图像哈希领域中许多算法只能处理灰度图像, 为扩大算法适用范围, 同时提高图像哈希算法的性能与旋转攻击的鲁棒性, 提出基于四元数拉盖尔矩和三维能量结构的哈希算法。首先对输入的彩色图像进行预处理与多尺度融合处理, 将融合图像提取的拉盖尔矩系数作为图像的全局特征, 同时在 YCbCr 颜色空间中利用融合图像的能量信息建立模型, 选取三维模型中不同视角下的能量峰值和谷值点连线与水平面的夹角作为局部结构特征; 然后根据特定点与三维模型各条等高线上近点和远点的位置提取具有旋转不变性的特征; 最后结合全局特征和三维结构特征, 量化并加密生成哈希序列。实验结果表明, 该算法在鲁棒性与区别性之间有更好的平衡, 受试者工作特征曲线错误接受率为 0 时的正确接受率达到 0.999 2。当哈希序列长度为 120 bit 时具备最优的紧凑性, 平均计算时间为 0.097 9 s。在拷贝检测实验中, 该算法进行多次抽取实验的平均查全率和查准率均在 95.83% 以上。

**关键词:** 四元数拉盖尔矩; 图像哈希; 图像能量; 三维结构特征; 拷贝检测

**源代码链接:** <https://gitee.com/suepbaijie/lagueree.git>

**中图分类号:** TP391

**文献标志码:** A

**DOI:** 10.19678/j.issn.1000-3428.0066413

## Image Hash Combining Quaternion Laguerre Moments and Three-Dimensional Structures

BAI Jie, ZHAO Yan\*

(College of Electronics and Information Engineering, Shanghai University of Electric Power, Shanghai 200090, China)

**【Abstract】** Currently, several algorithms used in image hashing can only handle grayscale images. In this study, a hashing algorithm based on quaternion Laguerre moments and a three-dimensional energy structure is proposed to improve the range of application and performance of the image hashing algorithm, particularly its robustness against rotation attacks. First, the input color image is preprocessed, multiscale fusion is performed, and the Laguerre moment coefficients are extracted from the fused image as the global features of the image. The energy information of the fused image is used to establish a model in the YCbCr color space, and the angle between the energy peak and valley points connected with the horizontal plane at different viewpoints in the three-dimensional (3D) model is selected as the local structure feature. The features with rotational invariance are extracted by the positions of the near and far points on the specific points and each contour of the 3D model. Finally, the global and 3D structural features are combined, quantized, and encrypted to generate hash sequences. Finally, the global and 3D structural features are combined to quantify and encrypt the generated hash sequences. The results show that the subject operating characteristic curve exhibits a correct acceptance rate of 0.999 2 when the error reception rate is 0. A Hash sequence length of 120 bit possesses optimal compactness, and the average computation time reaches 0.097 9 s. In copy detection experiments, the algorithm performs multiple extraction experiments with an average check-all rate, and the average check rate and accuracy rate of the algorithm for multiple extraction experiments are higher than 95.83%.

**【Key words】** quaternion Laguerre moments; image Hash; image energy; three-dimensional structural features; copy detection

## 0 引言

随着网络媒体的日益发展, 用户可以将更多的视频等包含大量图像的信息传入互联网。这些图像

通过 photoshop 等图像处理软件进行一定操作就可以产生诸如改变亮度、添加水印、马赛克等拷贝版本。若有人恶意篡改图像则可能带来严重后果, 因而正确区分这些图像至关重要。图像哈希技术的出

收稿日期: 2022-12-01 修回日期: 2023-01-15

基金项目: 国家自然科学基金(61802250); 上海市科委部分地方院校能力建设项目(20020500700)。

通信作者 E-mail: \*yanzhao79@hotmail.com

现为这类问题带来了解决方法。图像哈希是指将一副图像定向映射为一串字符或数字序列。哈希算法需要具备具有鲁棒性、区别性、安全性、紧凑性等基本特性,在图像分类、检索等领域中被广泛应用。

1) 基于变换的图像哈希。早期哈希主要通过图像变换来提取特征,如离散傅里叶变换(DFT)、离散余弦变换(DCT)、小波变换、对数极坐标变换等。文献[1]结合对数极坐标变换和离散傅里叶变换对图像提取旋转不变特征矩阵来构建哈希序列。文献[2]将 DFT 变换和 Radon 变换结合构建哈希,通过量化 DFT 变换系数和提取变换后的特征得到哈希。以上两种算法在对抗攻击的鲁棒性方面效果有限,且适用攻击操作种类较少。文献[3]考虑到对数极坐标变换具有良好的旋转不变性,结合 LT 变换和四元数傅里叶变换,将经过 LT 变换的次级图像提取 QDFT 幅度矩阵完成哈希算法,该算法鲁棒性较强但可对抗操作不多。文献[4]将图像进行非重叠分割,提取每个图像块的 DCT 系数并量化得到鲁棒的哈希序列。基于同样的非重叠分割思想,文献[5]使用 canny 算子计算图像块的特征点,选取代表性强的特征块和位置信息与 DCT 系数连接,经过主成分分析(PCA)得到哈希序列,这两种算法可对抗的攻击操作种类有所增加,但旋转攻击抗性都较弱。文献[6]使用 Radon 变换提取图像系数并进行一维 DCT,通过低频系数统计特征得到哈希序列,该算法的旋转抗性较优。文献[7]利用多级离散小波变换(DWT)提取特征构造哈希序列。文献[8]在小波变换基础上结合图像的特征点构造哈希序列,增强了算法的区别性。文献[9]提出将对数极坐标变换和 Gabor 变换相结合提取图像特征,该算法在区分性良好的基础上还具有较强的紧凑性。早期的图像哈希算法结构较为简单,主要聚焦于提高算法的鲁棒性和区别性,但对计算效率、攻击容量有所欠缺。因此,在保证算法鲁棒性的前提下,为了提高算法的处理效率,序列紧凑性也成为优化的目标。

2) 基于数据降维的图像哈希。文献[10]将局部线性嵌入用于哈希序列的压缩,降低了算法的计算代价。随后文献[11]提出使用图像的最大内接圆进行等面积环形分割将图像重组为列,再进行非负矩阵分解提取特征,该算法对旋转攻击的鲁棒性良好,但由于使用环分割导致图像边缘处的敏感性不足。文献[12]提取图像块纹理特征并压缩。文献[13]将小波分解和 CS-LBP 相结合,该算法性能良好,且具备了一定的拷贝检测能力。文献[14]提出利用

WeberLBP 提取图像的纹理特征,并与图像的颜色特征相结合构建哈希算法,使用 PCA 压缩数据使得算法具有不错的紧凑性。文献[15]在 Lab 颜色空间中结合环形分区和向量不变矩,提出了具备区别性和强鲁棒性的算法。

3) 基于颜色特征和图像矩的图像哈希。随着研究的深入,简单的变换方式和单一的特征构成已经不能满足对算法性能的要求。文献[16]将 RGB 彩色图像转换到 HSI 和 YCbCr 颜色空间,计算各分量均值和方差,提取局部颜色特征。文献[17]以图像颜色特征作为出发点,提取彩色图像的对立色并分块提取颜色特征,再对亮度图像进行四叉树分解得到结构特征,联合两种特征组成哈希序列。该算法具有强鲁棒性并且可以进行篡改定位。文献[18]使用图像的 Zernike 矩构成哈希序列,结合保角变换,提取图像的角信息,提高算法鲁棒性且可以进行篡改定位。

此外,文献[19]利用四元数奇异值分解计算出拼接图像的噪声信息,增强了图像的特征表达。文献[20]使用空洞卷积提取不同尺度的图像特征,将多尺度特征融合后作为网络的输入,有效增强了算法的检测性能。这些针对图像特征的提取和表达方面的改进值得学习借鉴。

目前的哈希算法在保证鲁棒性的前提下各具优势,但仍存在一些短板,如针对灰度图像的方案整体鲁棒性较强但不适用于彩色图像,可对抗的攻击种类较少。本文提出一种可以直接处理彩色图像并且在鲁棒性、区别性和序列紧凑性上具有一定优势的算法。由于拉盖尔矩对多数篡改操作具有鲁棒性,将其与四元数理论相结合拓展到图像哈希中,直接提取 RGB 图像的特征,得到的全局特征具有较好的鲁棒性。对预处理后的图像提取能量,在三维空间视角下提取能量的结构特征,在保证算法鲁棒性的同时提高区别性,增强对旋转攻击的鲁棒性。

## 1 本文图像哈希算法

### 1.1 算法框架

图 1 是结合四元数拉盖尔矩和三维结构特征的哈希算法流程,主要由预处理、特征提取、哈希生成组成(彩色效果见《计算机工程》官网 HTML 版,下同)。图像预处理是指对图像进行双线性插值得到二次图像,再对二次图像进行多尺度融合分别得到高斯融合图像和拉普拉斯融合图像。特征提取包括全局特征和三维能量结构特征,全局特征是指提取图像的四元数拉盖尔矩系数,三维能量结构特征是

指利用亮度图像的能量和非重叠分割后各子块的能量,在三维空间下提取结构特征。哈希生成将所得各部分特征量化组合得到中间哈希,经过密钥加密生成哈希序列。

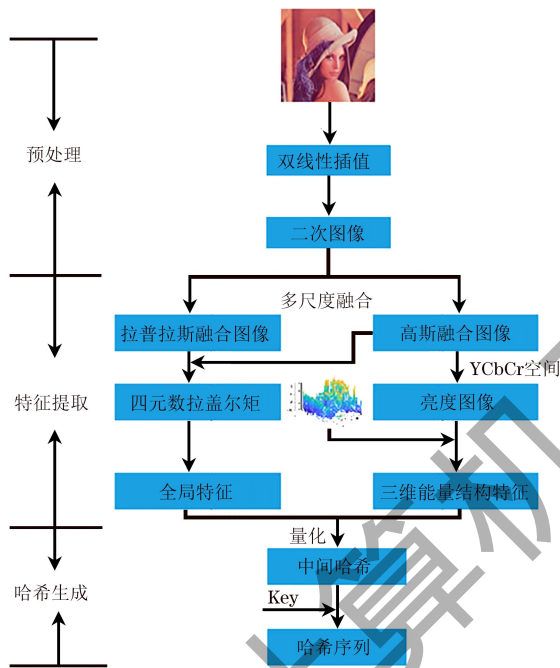


图 1 哈希算法流程

Fig.1 Procedure of Hash algorithm

## 1.2 图像预处理

用双线性插值统一输入图像尺寸为  $N \times N$ , 以处理不同大小的输入图像。该操作可以略微提升缩放攻击抗性,对统一尺寸图像  $I_0$  进行下采样。下采样的模糊图像记为  $I_1, I_2, I_3, I_4$ 。再用圆盘均值滤波器进行多级滤波。拉普拉斯金字塔表达式如式(1)所示:

$$\begin{cases} L_1 = I_0 - I_1 \\ L_2 = I_1 - I_2 \\ L_3 = I_2 - I_3 \\ L_4 = I_3 - I_4 \end{cases} \quad (1)$$

图像金字塔包含不同尺度下图像的信息。 $I_0, I_1, I_2, I_3$  为四级高斯金字塔,  $L_1, L_2, L_3, L_4$  为四级拉普拉斯金字塔,将图像金字塔进行融合可以发挥增强图像的作用。通过多尺度融合得到二次图像的高斯融合图像  $I_M$  和拉普拉斯融合图像  $L_M$ , 如式(2)所示:

$$\begin{cases} I_M = \frac{1}{4} \sum_{i=0}^3 I_i^* \\ L_M = \frac{1}{4} \sum_{i=1}^4 L_i^* \end{cases} \quad (2)$$

其中:  $I_i^*, L_i^*$  分别为  $I_i, L_i$  的尺度归一化图像。

## 1.3 全局特征提取

文献[21]提出在笛卡儿坐标系中灰度图像的广义拉盖尔矩,广义拉盖尔多项式可以由式(3)定义:

$$\bar{L}_n^{(\alpha, \lambda)}(x) = L_n^{(\alpha, \lambda)}(x) \sqrt{\frac{\omega^{(\alpha, \lambda)}(x)}{\gamma_n^{(\alpha, \lambda)}}} \quad (3)$$

其中:  $\alpha$  为平移参数,主要影响不同阶多项式的幅值和  $x$  轴的 0 点数量;  $\lambda$  为分数阶参数;  $\omega^{(\alpha, \lambda)}(x) = \lambda x^{(\alpha+1)\lambda-1} \exp(-x^\lambda)$ ;  $\gamma_n^{(\alpha, \lambda)} = \frac{\Gamma(n+\alpha+1)}{n!}$ 。

由式(3)可得,广义拉盖尔矩的定义如式(4)所示:

$$\text{FrS}_{nm}^{(\alpha, \lambda)} = \omega \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f^{\text{gray}}(i, j) \bar{L}_n^{(\alpha, \lambda)}(x_i) \bar{L}_m^{(\alpha, \lambda)}(y_j) \quad (4)$$

其中:  $f^{\text{gray}}(i, j)$  表示一幅  $N \times N$  的灰度图像。

为了方便计算,将原始图像矩阵映射到一个  $[0, K] \times [0, K]$  的方形区域,即:  $K > 0, \omega = \left(\frac{K}{N}\right)^2$ ,

$$x_i = \frac{iK}{N}, y_j = \frac{jK}{N}, i, j = 0, 1, \dots, N-1, n, m = 0, 1, \dots$$

为矩的阶数。通过调整  $\alpha$  的取值可以实现对图像的局部区域进行特征提取,本文需要提取图像的全部信息,故取  $\alpha=1$ 。调整  $\lambda$  的取值可以将整数阶的多项式扩展到实数阶,从而改变多项式在  $x$  轴上 0 点分布的紧凑程度,为了方便计算取  $\lambda=1$ 。所以,一幅 RGB 彩色图像在笛卡儿坐标系的右侧四元数拉盖尔矩定义如式(5)所示:

$$\text{QFrS}_{nm} = \omega \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \bar{L}_n(x_i) \bar{L}_m(y_j) f^{\text{rgb}}(i, j) \mu \quad (5)$$

其中:  $\mu = (i+j+k)/\sqrt{3}$  为单位纯虚数四元数。计算融合图像的四元数拉盖尔矩,提取其前  $n$  阶系数,矩系数在  $n$  取不同值时如表 1 所示。

由于两种融合图像求得的四元数拉盖尔矩系数是复数形式,因此对其求模并排列为行向量,则  $I_M$

表 1  $L_{n,m}$  系数

Table 1  $L_{n,m}$  coefficients

阶数 $n$	$L_{n,m}$ 矩	数量/个
1	$L_{0,0}$	1
2	$L_{1,0}, L_{1,1}$	3
3	$L_{2,0}, L_{2,1}, L_{2,2}$	6
4	$L_{3,0}, L_{3,1}, L_{3,2}, L_{3,3}$	10
5	$L_{4,0}, L_{4,1}, L_{4,2}, L_{4,3}, L_{4,4}$	15
6	$L_{5,0}, L_{5,1}, L_{5,2}, L_{5,3}, L_{5,4}, L_{5,5}$	21

的四元数拉盖尔矩系数为  $G = [g_1, g_2, \dots, g_{n(n+1)/2}]$ ,  $L_M$  的四元数拉盖尔矩系数为  $J = [j_1, j_2, \dots, j_{n(n+1)/2}]$ 。按式(6)分别对融合图像的矩系数进行二值化,  $G_M$  和  $J_M$  为  $G$  和  $J$  的均值,  $G_1$  和  $J_1$  为二值化向量,  $i = 1, 2, \dots, n(n+1)/2$ 。

$$\begin{aligned} G_1(i) &= \begin{cases} 1, G(i) > G_M \\ 0, \text{其他} \end{cases} \\ J_1(i) &= \begin{cases} 1, J(i) > J_M \\ 0, \text{其他} \end{cases} \end{aligned} \quad (6)$$

连接融合图像的矩系数得到全局特征  $H_Q = [G_1, J_1]$ , 长度为  $L_1 = n(n+1)$ 。

#### 1.4 三维能量结构特征提取

文献[22]提出图像在经过常见的攻击操作前后, 其能量并不会发生大的波动, 所以可以认为图像能量对于常见攻击操作有较好的鲁棒性。算法在 YCbCr 颜色空间下提取高斯融合图像的亮度图像  $Y$ , 对于尺寸为  $N \times N$  的亮度图像, 它的能量  $E$  可以表示为:

$$E(Y) = \sum_{i=1}^N \sum_{j=1}^N y_{i,j}^2 = \text{trace}(Y^T Y) \quad (7)$$

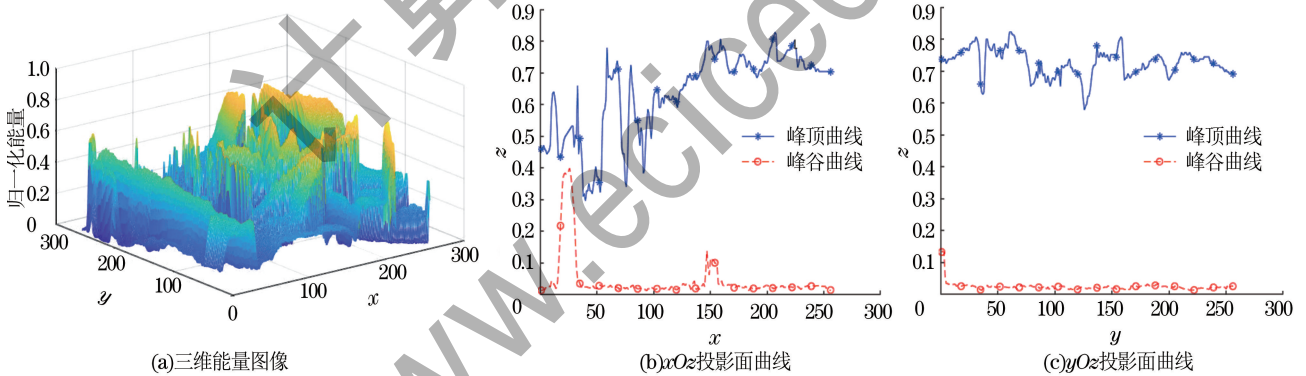


图 2 三维能量模型和峰谷值曲线

Fig.2 Three-dimensional energy model and peak and valley value curves

按照式(8)获得  $N_E$  在  $xOz$  投影面下的峰顶曲线  $M_1$  和峰谷曲线  $M_2$ , 同理可以得到  $N_E$  在  $yOz$  投影面下的峰顶曲线  $M_3$  和峰谷曲线  $M_4$ 。

$$M_1 = \max(N_E, 1), M_2 = \min(N_E, 1) \quad (8)$$

其中:  $\min(\cdot, 1)$  和  $\max(\cdot, 1)$  表示按行取  $N_E$  最小值和最大值。在不同投影面的峰谷值曲线上等间距取出 16 对峰谷值点, 记  $xOz$  投影面取出的峰值点集为  $M_{1P}$ , 谷值点集为  $M_{2V}$ ,  $yOz$  投影面取出的峰值点集为  $M_{3P}$ , 谷值点集为  $M_{4V}$ 。提取  $xOz$  投影面峰值点集  $M_{1P}$  在  $xOy$  平面上的位置信息矩阵  $P = [P_1, P_2, \dots, P_{16}]$  与  $yOz$  投影面峰值点集  $M_{3P}$  在  $xOy$  平面上的位置信息矩阵  $U = [U_1, U_2, \dots, U_{16}]$ 。同理,  $xOz$  投影面谷值点集  $M_{2V}$  在  $xOy$  平面上的位置信息矩阵

其中:  $\text{trace}(\cdot)$  表示矩阵的迹;  $y_{i,j}$  表示亮度图像  $Y$  的像素值。本文选择在三维空间内进行特征提取, 图像的三维能量结构特征由两部分组成。第一部分先计算亮度图像  $Y$  的能量矩阵  $N_E$ , 以  $N_E$  的行位置为  $x$  轴、列位置为  $y$  轴、能量值为  $z$  轴构建三维模型, 在  $xOz$  和  $yOz$  两个投影面上画出峰顶曲线和峰谷曲线, 求取峰值点与谷值点连线的夹角余弦值。第二部分以  $b \times b$  的子块大小非重叠分割亮度图像, 计算各子块能量值得到亮度图像的分块能量矩阵  $N_Y$ , 以  $N_Y$  的行位置为  $x$  轴、列位置为  $y$  轴、能量值为  $z$  轴构建三维模型并画出等高线, 求出特定点到等高线上最近点和最远点连线的余弦值, 从而充分提取三维结构特征。

本文算法三维特征在构造时主要考虑: 1) 使用图像能量构造三维模型; 2) 通过提取峰谷值点和分重叠分割使哈希序列更紧凑; 3) 加入等高线特征增强抵抗旋转攻击的能力。

三维能量图像和相应投影面的峰谷值曲线如图 2 所示。

$Q = [Q_1, Q_2, \dots, Q_{16}]$ ,  $yOz$  投影面谷值点集  $M_{4V}$  在  $xOy$  平面上的位置信息矩阵  $V = [V_1, V_2, \dots, V_{16}]$ 。将  $xOz$  投影面上的峰值点集按照序数从大到小依次与谷值点集序数从小到大的顺序相连, 形成线段集  $S = [P_1 Q_{16}, P_2 Q_{15}, \dots, P_{16} Q_1]$ , 共计 16 条线段。同理, 取得  $yOz$  投影面上的峰值点与谷值点的线段集  $K = [U_1 V_{16}, U_2 V_{15}, \dots, U_{16} V_1]$ 。记  $xOy$  平面为  $F$ , 求出这些线段与  $xOy$  平面夹角的余弦值, 将一系列余弦值排列成行向量并按式(9)进行二值化。

$$\begin{aligned} H_1(i) &= \begin{cases} 1, \cos(S_i, F) > S_M \\ 0, \text{其他} \end{cases} \\ H_2(i) &= \begin{cases} 1, \cos(K_i, F) > K_M \\ 0, \text{其他} \end{cases} \end{aligned} \quad (9)$$

其中： $S_M$  表示  $xOz$  平面线段集与水平面  $F$  夹角余弦值的平均值； $K_M$  表示  $yOz$  平面线段集与水平面  $F$  夹角余弦值的平均值， $i = 1, 2, \dots, 16$ ； $H_1, H_2$  长度都为 16。以上是通过图像在三维能量模型上不同投影面的特征点得到局部结构特征。

为了进一步提高算法对于旋转攻击的鲁棒性，接下来对图像进行抗旋转特征提取。使用图像子块能量值作为等高线特征的  $z$  轴主要原因如下：1) 图像子块的能量值对常见的攻击操作同样具有良好的鲁棒性；2) 分块处理可以在不影响算法鲁棒性的前提下有效降低数据维度，提高计算效率；3) 为了提高对于旋转攻击的鲁棒性，选择在三维结构上寻找等高线提取抗旋转的结构特征，图像分块处理可以使算法以微弱降低 ROC 曲线

表现力为代价，明显提升对旋转攻击的鲁棒性。具体步骤如下：算法利用亮度图像的分块能量矩阵  $N_V$  的信息，以行位置信息为  $x$  轴、列位置信息为  $y$  轴、子块能量值为  $z$  轴构造三维模型。在三维结构上画出一系列等高线，选定  $xOy$  平面的中心点  $O$  作为原点，找到原点在每一条等高线上距离最近和最远的两点。这样做的优点是当图像受到任意角度的旋转攻击，分块能量矩阵的三维模型会随之发生旋转，但等高线只会发生水平旋转，所选取的原点到等高线上任意一点的连线与水平面的夹角在旋转前后不会发生改变。所以选取这样一组在任意角度旋转攻击下保持不变的特征有助于提高算法的鲁棒性。分块能量矩阵  $N_V$  的三维模型等高线和等高线在  $xOy$  平面的投影示意图如图 3 所示。

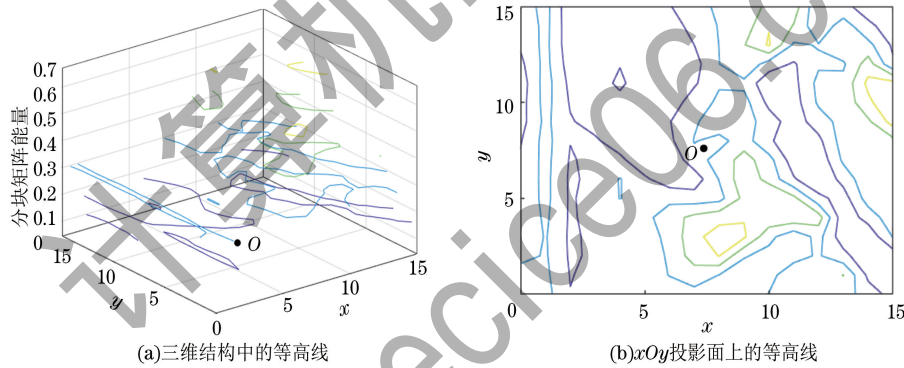


图 3 等高线模型示意图

Fig.3 Schematic diagram of the contour model

为便于观察，图 3 显示了取 4 条等高线的三维结构示意图，且标注出了中心点  $O$  所在位置。求解原点到等高线近点远点连线与水平面  $F$  的余弦值，记为近点余弦值集  $D_{\min} = [D_1, D_2, \dots, D_n]$ ，远点余弦值集  $D_{\max} = [d_1, d_2, \dots, d_n]$ ，将其按式(10)进行二值化：

$$\begin{aligned} H_3(i) &= \begin{cases} 1, D_{\min}(i) > D_M \\ 0, \text{其他} \end{cases} \\ H_4(i) &= \begin{cases} 1, D_{\max}(i) > D_M \\ 0, \text{其他} \end{cases} \end{aligned} \quad (10)$$

其中： $D_M$  为所有连线与水平面夹角余弦值的平均值； $i = 1, 2, \dots, n$ ； $H_3, H_4$  长度都为  $n$ 。通过联合投影面峰谷值点的局部结构特征和分块能量矩阵的等高线特征，得到图像的三维能量结构特征  $H_E = [H_1, H_2, H_3, H_4]$ ， $L_2 = 2n + 32$  表示  $H_E$  的长度。

### 1.5 哈希生成

连接  $H_Q$  和  $H_E$  得到中间哈希  $H_{\text{mid}} = [H_Q,$

$H_E]$ ，本文算法得到的哈希序列长度为  $L = L_1 + L_2 = n(n+1) + 2n + 32$ 。出于安全性考虑，用随机生成的 1 000 个伪随机数列  $w$  重排中间哈希  $H_{\text{mid}}$  得到最终哈希序列  $h$ ，如式(11)所示：

$$h(i) = H_{\text{mid}}(w[i]) \quad (11)$$

其中： $w(i)$  表示伪随机数序列  $w$  的第  $i$  个数。本文算法得到二进制哈希序列，选择汉明距离作为图像相似度判断依据，当两幅图像的汉明距离小于阈值  $T$  时判断为相似图像，反之为不同图像。汉明距离  $D$  的计算公式如式(12)所示：

$$D = \sum_{i=1}^L (h_1(i) \oplus h_2(i)) \quad (12)$$

其中： $h_1(i)$  和  $h_2(i)$  是  $h$  中的第  $i$  个元素； $\oplus$  为异或运算； $h_1, h_2$  表示进行判断的两幅图像。

## 2 实验结果与分析

本节主要进行包括鲁棒性实验、区别性实验、参数分析、不同算法对比。实验中使用的参数如下：图像归一化尺寸  $N = 256$ ，图像子块大小  $b =$

16, 阶数  $n=8$ , 故算法的哈希序列长度为  $L=120$  bit。所有的实验均通过 MATLABR2019a 平台进行仿真。

## 2.1 鲁棒性实验

鲁棒性实验中测试样本由图 4 所示的 20 幅彩色图像组成。



图 4 鲁棒性实验图像

Fig. 4 Images of robustness experiments

按照如表 2 所示的操作攻击每幅测试图像, 单个测试图有 66 个相似版本, 将测试图像与相似图像配对, 计算相似图像对的哈希距离。

表 2 鲁棒性实验攻击参数

Table 2 Attack parameters for robustness experiments

图像处理	参数	参数值	数量
JPEG 压缩	质量因子	0.3, 0.4, ..., 1.0	8
对比度调整	级别	-20, ..., 20	4
伽马校正	$\gamma$ 值	0.75, ..., 1.25	4
亮度调整	级别	-20, ..., 20	4
椒盐噪声	级别	0.002, ..., 0.010	5
图像缩放	比例	0.6, 0.8, ..., 1.8	6
水印嵌入	透明度	0.3, 0.4, ..., 0.8	6
高斯低通滤波	标准差	0.1, ..., 1.0	10
乘性噪声	方差	0.002, ..., 0.010	5
均值滤波	模板大小	$3 \times 3, 5 \times 5, \dots, 9 \times 9$	4
旋转	角度/ $^\circ$	0.2, 0.4, ..., 1.0	5
高斯噪声	均值	0.002, ..., 0.010	5

表 3 是 20 幅测试图像经过以上 12 类攻击操作得到的相似图像与原图像的哈希距离统计。由表 3 可以看出, 相似图像与原图像的哈希距离都落在 0~18 区间内, 哈希距离均值都不超过 12。说明本文算法对常见的攻击操作具有良好的鲁棒性, 并且当阈值设置到 18 以上时, 理论上可以将全部的相似图像识别出来。

表 3 哈希距离表

Table 3 Hash distance table

图像处理	最小值	最大值	均值	标准差
JPEG 压缩	0	10	3.85	3.11
对比度调整	1	6	3.10	2.71
伽马校正	2	15	6.85	3.79
亮度调整	1	8	4.15	3.47
椒盐噪声	2	12	6.76	3.52
图像缩放	0	19	6.53	5.07
乘性噪声	1	12	6.88	4.02
高斯低通滤波	1	8	4.32	0.82
高斯噪声	5	16	10.40	2.84
均值滤波	3	18	11.25	4.80
旋转/ $^\circ$	5	14	11.16	7.23
水印嵌入	0	1	0.33	0.41

图 5 展示了图 4 中前 5 个测试样本在不同篡改攻击下的哈希距离分布, 可得算法对多种常见攻击的鲁棒性都较好。

## 2.2 区别性实验

区别性实验数据集共有 1 000 张各类图片, 其中 700 张来自 Ground Truth Database<sup>[23]</sup> 数据库, 300 张来自 VOC2007<sup>[24]</sup> 数据库。对区别性实验数据集的所有图像两两计算汉明距离, 得到  $C_{1000}^2 = 499\,500$  个不同图像对。选择表 2 中列举出的 11 类共 22 种篡改方式如表 4 所示。对数据集内的图像

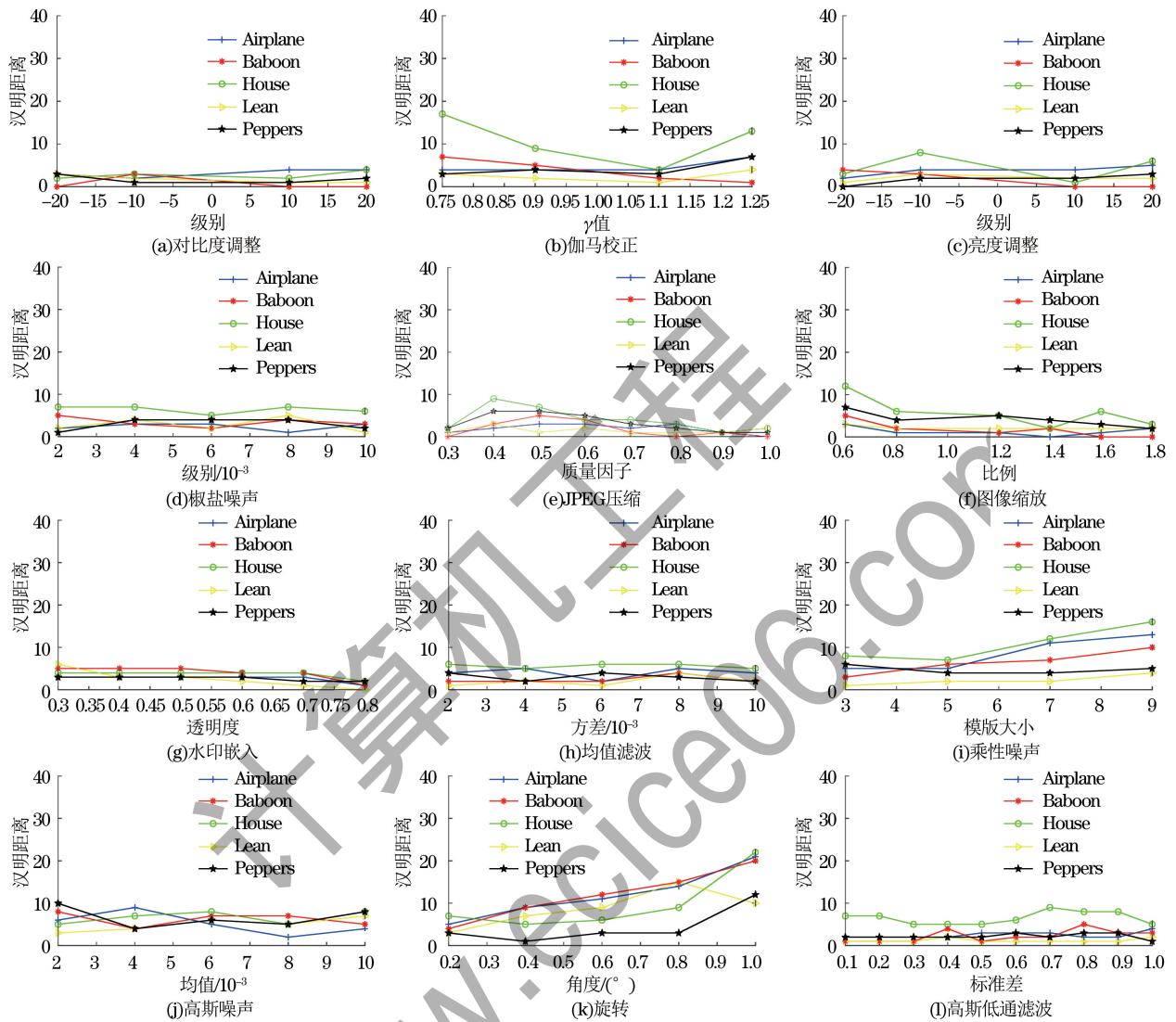


图 5 鲁棒性实验结果

Fig.5 Results of robustness experiments

进行篡改,一张原始图片能生成 22 幅相似图片,相似图像对总数为  $C_{23}^2 \times 1\,000 = 253\,000$ 。

表 4 区别性实验攻击参数

Table 4 Attack parameters of the differentiation experiment

图像处理	参数设置	参数值
JPEG 压缩	质量因子	0.4, ..., 0.8
对比度调整	级别	-20, ..., 20
伽马校正	伽马值	0.75, ..., 1.25
亮度调整	级别	-20, ..., 20
椒盐噪声	级别	0.002, ..., 0.006
图像缩放	比例	0.5, ..., 1.5
乘性噪声	方差	0.002, ..., 0.006
高斯低通滤波	标准差	0.1, 0.2, ..., 1.0
高斯噪声	均值	0.002, ..., 0.006
水印嵌入	透明度	0.3, ..., 0.8
均值滤波	模板大小	3×3, ..., 5×5

图 6 展示了不同图像对和相似图像对的哈希距离分布。其中,横坐标表示汉明距离,纵坐标表示该距离下相似或不同图像对数目。由图 6 可以看出,相似图像对汉明距离区间为 0~45,不同图像对汉明距离区间为 22~118,即在汉明距离为 22~45

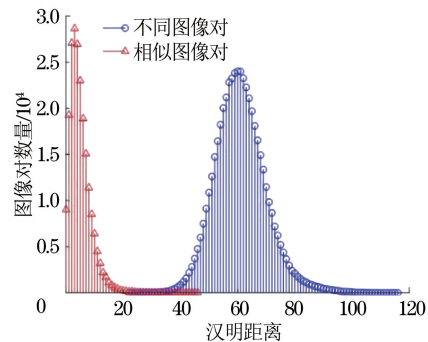


图 6 汉明距离分布

Fig.6 Hamming distance distribution

时,相似图像对和不同图像对发生了重叠。算法选取碰撞率( $P_C$ )和检错率( $P_E$ )指标衡量区别性,根据式(13)计算确定合适的阈值区分相似和不同图像对。

$$\begin{aligned} P_C &= N_C / N_D \\ P_E &= N_E / N_S \end{aligned} \quad (13)$$

其中: $N_C$ 表示不同图像误判相似图像数; $N_D$ 为不同图像数; $N_E$ 表示相似图像误判不同图像数; $N_S$ 为相似图像数。计算结果如表 5 所示,由图 6 和表 5 可知,当阈值  $T=34$  时,碰撞率为  $7.80 \times 10^{-5}$ ,检错率为  $2.42 \times 10^{-5}$ ,综合考虑碰撞率和检错率,选择  $T=34$  为阈值。

表 5 各阈值检错率和碰撞率

Table 5 Error detection rate and collision rate for each thresholds

阈值	$P_C$	$P_E$
22	0	$7.67 \times 10^{-4}$
26	$1.20 \times 10^{-5}$	$2.55 \times 10^{-4}$
30	$3.60 \times 10^{-5}$	$9.19 \times 10^{-5}$
34	$7.80 \times 10^{-5}$	$2.42 \times 10^{-5}$
38	$2.41 \times 10^{-4}$	$5.33 \times 10^{-6}$
45	$5.26 \times 10^{-4}$	0

## 2.3 参数分析

### 2.3.1 阶数 $n$ 和分块大小 $b$ 对算法的影响

实验在除阶数和分块大小以外的条件固定情况下,取  $n=[6,7,8,9]$ , $b=[8,16]$ ,两两组合分别进行实验。通过 ROC<sup>[25]</sup> 曲线衡量选取阶数和分块大小的不同取值时算法的综合性能如图 7 所示。参数分析实验的数据集与前文区别性实验相同。通过错误接受率( $P_{FPR}$ )和正确接受率( $P_{TPR}$ )指标衡量区别性,计算公式如式(14)所示:

$$\begin{aligned} P_{FPR} &= n_1 / N_1 \\ P_{TPR} &= n_2 / N_2 \end{aligned} \quad (14)$$

其中: $n_1$ 指错判相似图像对数; $n_2$ 为正确判定相似图像对数; $N_1$ 和  $N_2$ 分别为不同图像对和相似图像对的总数。

图 7 显示,当  $n=8$ 、 $b=16$  时,算法的 ROC 曲线表现力最佳,且当  $b$  的取值固定为 8 或 16 时,曲线表现力都随着  $n$  由 6 增加到 8 持续提高,但  $n=9$  时哈希序列长度达到 140,等高线特征长度达到 18,一定程度上影响了算法的综合表现力,故综合考虑当  $n=8$ 、 $b=16$  时为最佳选择。

### 2.3.2 多尺度融合处理的效果

由于所提算法在预处理的过程中为了强化对

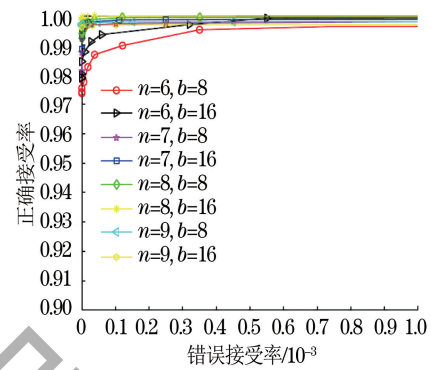


图 7 各分块大小和阶数下的 ROC 曲线

Fig. 7 ROC curves for each block size and order

图像信息的表达,对插值以后的二次图像进行了多尺度融合的处理,因此在所有参数都相同的情况下,将经过多尺度融合的算法和直接使用二次图像进行特征提取并计算哈希的算法进行比较。ROC 曲线对比如图 8 所示,说明算法经过多尺度融合后,区别性有所提高。

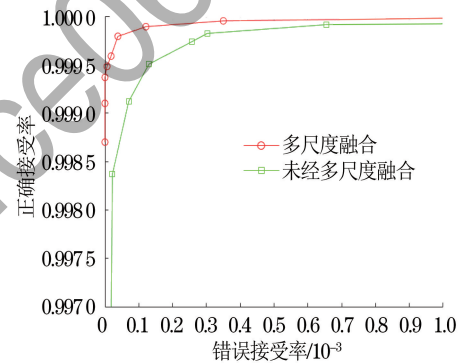


图 8 多尺度融合前后的 ROC 曲线

Fig. 8 ROC curves before and after multi-scale fusion

### 2.3.3 等高线特征对旋转抗性的影响

为了验证等高线特征对算法在旋转攻击鲁棒性方面的影响,在其他参数相同的情况下,不添加等高线特征的算法对旋转攻击的鲁棒性和加入等高线特征的旋转攻击鲁棒性如图 9 所示。

图 9(a)是无等高线特征的旋转攻击鲁棒性图,哈希序列长度为 104,可以看出,在  $0.2^\circ \sim 1.0^\circ$  的旋转攻击下,哈希距离最高为 22,且上升趋势明显。图 9(b)为加入等高线特征的旋转攻击鲁棒性,哈希序列长度为 120,由图 9(b)显示在  $0.2^\circ \sim 1.0^\circ$  的旋转攻击下,哈希距离最大为 14,图 9(c)显示在  $1.0^\circ \sim 5.0^\circ$  旋转攻击下,哈希距离最大为 26。可以看出,在哈希序列长度增加的同时,原始图像与经过旋转攻击的图像之间的哈希距离反而有所降低,证明等高线特征的加入有效提高了算法对于旋转攻击的鲁棒性和综合表现力。

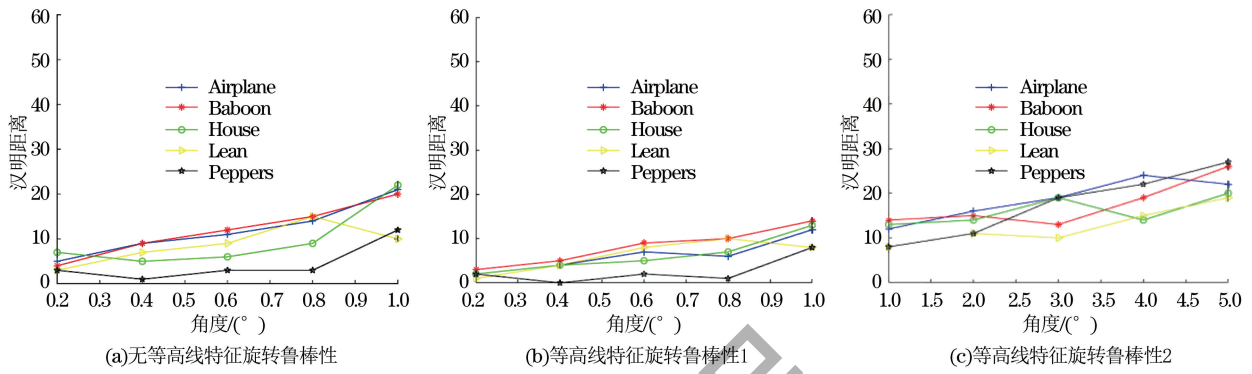


图 9 加入等高线特征前后的鲁棒性对比

Fig. 9 Robustness comparison before and after adding contour features

2.4 图像分类性能对比

为考量本文算法的图像分类性能,将本文算法与文献[12,26-28]算法的 ROC 曲线性能进行对比。为保证公平性,本节参与对比的各个算法使用同样的数据集并且在相同平台运行。图 10 展示了不同算法之间 ROC 曲线对比。

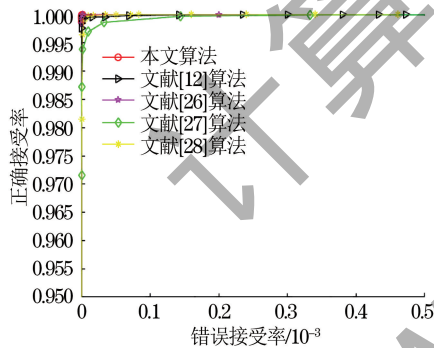


图 10 不同算法的 ROC 曲线对比

Fig. 10 ROC curve comparison of different algorithms

当  $P_{FPR} = 0$  时,本文算法的  $P_{TPR}$  为 0.999 2。参与对比的算法<sup>[12, 26-28]</sup>  $P_{TPR}$  分别为 0.998 5、0.999 1、0.924 5 和 0.996 4。当  $P_{TPR}$  接近于 1 时,本文算法的  $P_{FPR}$  为 0.000 6,文献[12, 26-28]算法的  $P_{FPR}$  分别为 0.112 1、0.002 0、0.761 1 和 0.060 4。可以看出,当  $P_{FPR} = 0$ 、 $P_{TPR} = 1$  时,本文算法的表现都优于对比算法。在算法的哈希生成效率和储存代价方面,计算了同样外部参数情况下本文算法与各对比算法的总运行时间和哈希长度。将 1 000 张输入图像产生哈希序列的总时间记录下来并除以 1 000 得到各个算法的平均哈希生成时间。表 6 综合记录了本文算法与对比算法的性能对比。通过表 6 可以看出,本文算法的哈希序列长度为 120 bit,在所有对比算法中紧凑性最好,运算速度快。图 10 展示出本文算法的 ROC 曲线最接近图像的左上角,说明本文算法的综合性能较好。

表 6 不同算法的性能对比

Table 6 Performance comparison of different algorithms

算法	哈希长度	平均时间/ s	$P_{FPR} = 0$ 时 最佳 $P_{TPR}$	$P_{TPR} = 1$ 时 最佳 $P_{FPR}$
本文算法	120 bit	0.097 9	0.999 2	0.001 1
文献[12]算法	64 decimal digits	0.221 3	0.998 5	0.112 1
文献[26]算法	412 bit	0.110 1	0.999 1	0.002 0
文献[27]算法	192 decimal digits	0.984 8	0.924 5	0.761 1
文献[28]算法	64 decimal digits	0.567 4	0.996 4	0.060 4

由于等高线特征通过分块能量矩阵得到,若图像归一化尺寸  $N$  发生变化,等高线特征就会随之变化而改变算法性能。为了展现本文算法在不同  $N$  值下的性能,选择  $N = 256$  和  $N = 512$ ,通过 ROC 曲线来衡量算法性能,如图 11 所示。可以看出,本算法在  $N = 512$  时性能发生了轻微下滑,这是由于虽然  $N$  取值的变化会使图像的分块能量矩阵发生变化,但等高线特征的长度并不会受到影响,从而将归一化尺寸波动带来的影响控制在可接受的范围

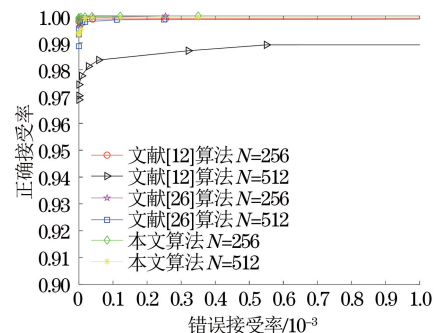


图 11 各算法在不同 N 值下 ROC 曲线对比

Fig. 11 Comparison of ROC curve of each algorithms under different N values

内。文献[12]算法性能下滑非常明显,这是由于  $N$  取值提高导致哈希序列过长,在区别性上的表现下降明显,从而影响了 ROC 曲线。文献[26]算法也有一定程度的削弱。由图易得,在  $N=256$  时本算法具有 ROC 曲线表现上的优势。

### 2.5 拷贝检测实验

图像检索在实际应用中面对的目标应该是互联网媒体中的各种图像,但在实验研究中通常使用已有的数据库,并对这些数据库中的图像进行攻击处理来验证算法的图像分类性能。拷贝检测实验则是在网络上随机下载图像,并对图像进行攻击处理。原始图像的来源不同可能导致分辨率、尺寸、颜色空间上的差异。这些差异在实验中会体现为算法可能在对某些数据库上能表现出较好的图像分类性能,但在拷贝检测实验中表现会有所降低。引入文献[29]提出的查全率( $P$ )和查准率( $R$ )作为衡量算法在拷贝检测方面性能的指标。随机下载 1 000 幅网络图像为拷贝检测原始图像库,其中抽取 100 幅作为查询图像库。对查询图像库的图片进行 28 种攻击操作,如表 7 所示。将得到的拷贝图像和原始图像库共计 3 800 幅作为实验图像库,对查询图像库进行实验。 $P$  和  $R$  的计算公式如式(15)所示:

$$\begin{aligned} P &= N_3 / N_4 \\ R &= N_3 / N_5 \end{aligned} \quad (15)$$

其中: $N_3$  是结果中正确查询拷贝图像数; $N_4$  是结果中拷贝图像数; $N_5$  是测试图像库中拷贝图像数。

表 7 拷贝检测攻击参数

Table 7 Attack parameters for copy detection

图像处理	参数设置	参数值
JPEG 压缩	质量因子	40, ..., 80
对比度调整	级别	-20, ..., 20
伽马校正	伽马值	0.75, ..., 1.25
亮度调整	级别	-20, ..., 20
椒盐噪声	级别	0.002, ..., 0.006
图像缩放	比例	0.8, ..., 1.6
乘性噪声	方差	0.002, ..., 0.006
3×3 高斯低通滤波	标准差	0.2, ..., 0.6
高斯噪声	均值	0.002, ..., 0.006
均值滤波	模板大小	3×3, ..., 5×5
旋转	角度/(°)	0.5, ..., 1.0
马赛克	方形大小	6, ..., 10
加字幕	字体大小	10, ..., 20
水印嵌入	透明度	0.3, ..., 0.8

表 8 是进行 5 次随机抽取后,拷贝图像与查询图像之间汉明距离在选择不同的阈值下算法的平均查全率和查准率。当得到的汉明距离小于所选阈值

时,认为该图像为拷贝图像,反之则认为图像是实验数据库中除拷贝图像之外的其他图像。由表 8 可以看出,当阈值为 30 时,算法能以 94.92% 的准确率将所有拷贝图像识别出来;当阈值取 38 时,查全率为 97.58%,查准率为 96.77%。

表 8 平均查全率和查准率

Table 8 Average completion rate and accuracy rate

阈值	查准率/%	查全率/%
30	94.92	100.00
32	95.87	99.86
34	96.36	99.33
36	96.61	98.67
38	96.77	97.58

同时为了增强查全率与查准率指标的稳定性,选取 1 000 幅图像中的 400 幅进行多次随机抽取实验得到的平均查全率与查准率如表 9 所示。

表 9 增加查询图像后的平均查全率和查准率

Table 9 Increase the average search completion rate and accuracy rate after querying images

阈值	查准率/%	查全率/%
30	92.79	100.00
32	94.17	98.91
34	94.86	97.73
36	95.41	97.15
38	95.83	96.62

通过拷贝检测实验与文献[12,22,26-27]算法进行比较,画出查全率和查准率曲线,结果如图 12 所示。可以看出,本文算法 P-R 曲线最靠近右上角,优于其他算法。结果表明,本文算法在拷贝检测应用方面具有较好的表现。

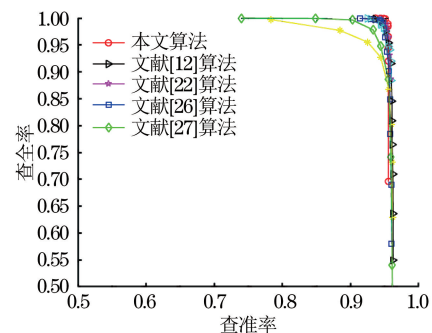


图 12 不同算法的 P-R 曲线对比

Fig.12 Comparison of P-R curve of different algorithms

### 3 结束语

本文提出一种结合四元数拉盖尔矩和三维能量结构的哈希算法。用四元数拉盖尔矩提取融合彩色图像的全局特征,对三维空间下高斯融和图像的能

量信息建立模型并提取结构特征,同时对非重叠分割后的能量信息构建三维模型,提取其中抗旋转的等高线特征,提高鲁棒性。实验结果表明,该算法的哈希长度紧凑,运算速度快,鲁棒性与区别性之间取得了较好的平衡。下一步将尝试对等高线特征进行拓展,使算法在保证性能的前提下增强对更大角度旋转攻击的鲁棒性。

#### 参考文献

- [1] OUYANG J L, COATRIEUX G, SHU H Z. Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform [J]. *Digital Signal Processing*, 2015, 41: 98-109.
- [2] LEI Y Q, WANG Y G, HUANG J W. Robust image hash in Radon transform domain for authentication [J]. *Signal Processing*, 2011, 26(6): 280-288.
- [3] TANG Z J, HUANG Z Q, ZHANG X Q, et al. Robust image hashing with multidimensional scaling [J]. *Signal Processing*, 2017, 137: 240-250.
- [4] TANG Z J, YANG F, HUANG L Y, et al. Robust image hashing with dominant DCT coefficients [J]. *Optik*, 2014, 125(18): 5102-5107.
- [5] QIN C, CHEN X Q, DONG J, et al. Perceptual image hashing with selective sampling for salient structure features [J]. *Displays*, 2016, 45: 26-37.
- [6] OU Y, RHEE K H. A key-dependent secure image hashing scheme by using Radon transform [C]//*Proceedings of International Symposium on Intelligent Signal Processing and Communication Systems*. Washington D. C., USA: IEEE Press, 2010: 595-598.
- [7] SHEN Q, ZHAO Y. Statistical feature hashing based on wavelet decomposition [J]. *Journal of Applied Sciences*, 2018, 36(2): 247-254.
- [8] VADLAMUDI L N, VADELLE R P V, DEVARA V. Robust image hashing using SIFT feature points and DWT approximation coefficients [J]. *ICT Express*, 2018, 4(3): 154-159.
- [9] LI X W, XIA X Z. Robust image hashing based on content structure diagram [J]. *Journal of Applied Sciences*, 2016, 34(6): 691-701.
- [10] TANG Z J, RUAN L L, QIN C, et al. Robust image hashing with embedding vector variance of LLE [J]. *Digital Signal Processing*, 2015, 43: 17-27.
- [11] TANG Z J, ZHANG X Q, ZHANG S C. Robust perceptual image hashing based on ring partition and NMF [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(3): 711-724.
- [12] DAVARZANI R, MOZAFFARI S, YAGHMAIE K. Perceptual image hashing using center-symmetric local binary patterns [J]. *Multimedia Tools and Applications*, 2016, 75(8): 4639-4667.
- [13] SHEN Q, ZHAO Y. Image hash algorithm for copy detection [J]. *Application Research of Computers*, 2019, 36(2): 611-614.
- [14] QIN C, HU Y C, YAO H, et al. Perceptual image hashing based on weber local binary pattern and color angle representation [J]. *IEEE Access*, 2019, 7: 45460-45471.
- [15] TANG Z J, ZHANG X Q, LI X X, et al. Robust image hashing with ring partition and invariant vector distance [J]. *IEEE Transactions on Information Forensics and Security*, 2015, 11(1): 200-214.
- [16] TANG Z J, ZHANG X Q, DAI X, et al. Robust image hash function using local color features [J]. *International Journal of Electronics and Communications*, 2013, 67(8): 717-722.
- [17] SHEN Q, ZHAO Y. Perceptual hashing for color image based on color opponent component and quadtree structure [J]. *Signal Processing*, 2020, 166: 107244.
- [18] ZHAO Y, WANG S Z. Perpetual image hashing based on conformal mapping and Zernike moments [J]. *Journal of Applied Sciences*, 2012, 30(1): 75-81.
- [19] 赵秀锋, 魏伟一, 陈金寿, 等. 基于自适应四元数奇异值分解的图像拼接检测 [J]. *计算机工程*, 2022, 48(4): 223-230, 239.
- [20] ZHAO X F, WEI W Y, CHEN J S, et al. Image splicing detection based on adaptive quaternion singular value decomposition [J]. *Computer Engineering*, 2022, 48(4): 223-230, 239. (in Chinese)
- [21] 赫晓慧, 宋定君, 李盼乐, 等. 融合多尺度特征的遥感影像道路提取方法 [J]. *计算机工程*, 2022, 48(8): 196-205.
- [22] HE X H, SONG D J, LI P L, et al. Remote sensing image road extraction method combined with multi-scale features [J]. *Computer Engineering*, 2022, 48(8): 196-205. (in Chinese)
- [23] PAN B J, LI Y H, ZHU H Q. Image description using radial associated laguerre moments [J]. *Journal of ICT Research and Applications*, 2015, 9(1): 1-19.
- [24] 李旭东. 基于图像能量和量化的公开数字水印算法 [J]. *中北大学学报(自然科学版)*, 2007, 28(5): 458-461.
- [25] LI X D. Publicdigital watermark algorithm based on image energy and quantization [J]. *Journal of North University of China(Natural Science Edition)*, 2007, 28(5): 458-461. (in Chinese)
- [26] University of Washington. Ground truth database [EB/OL]. [2022-10-30]. <http://www.cs.washington.edu/research/image-database/groundtruth/>.
- [27] REDMOD J. Pascal VOC dataset mirrir [EB/OL]. [2022-10-30]. <https://pjreddie.com/projects/pascal-voc-dataset-mirror/>.
- [28] FAWCETT T. An introduction to ROC analysis [J]. *Pattern Recognition Letters*, 2006, 27(8): 861-874.
- [29] MA L S, ZHAO Y. Image hashing algorithm based on QBFM moments and three-dimensional structure [J]. *Application Research of Computers*, 2022, 39(3): 949-955.
- [30] TANG Z J, ZHANG H Y, PUN C M, et al. Robust image hashing with visual attention model and invariant moments [J]. *IET Image Processing*, 2020, 14(5): 901-908.
- [31] TANG Z J, YU M Z, YAO H, et al. Robust image hashing with singular values of quaternion SVD [J]. *The Computer Journal*, 2021, 64(11): 1656-1671.
- [32] MÜLLER H, MÜLLER W, SQUIRE D M, et al. Performance evaluation in content-based image retrieval: overview and proposals [J]. *Pattern Recognition Letters*, 2001, 22(5): 593-601.