

面向仿射目标识别的几何与仿生融合特征提取方法

余伶俐¹, 易倩¹, 金鸣岳¹, 周开军²

(1. 中南大学自动化学院, 湖南长沙 410083; 2. 湖南工商大学智能工程与智能制造学院, 湖南长沙 410205)

摘要: 针对由于拍摄视角不同, 目标图像在水平或垂直方向发生拉长或压缩等仿射变换, 进而无法正确识别的问题, 本文设计了一种几何与仿生融合的特征提取方法. 首先, 对传统的角点和直线检测进行改进, 提出自适应Harris角点检测方法和去冗余的直线检测方法, 并将角点数、直线数和面积比向量作为几何特征. 然后, 采用生物启发变换算法提取图像的仿生启发特征, 该算法包括两个阶段, 每个阶段均需执行方向边缘检测和局部空间频率检测. 接着, 将输入图像的两种特征向量分别与标准数据库中的特征向量进行Pearson相关距离计算, 获得匹配得分. 最后, 在考虑不同数据库两种特征区分性强弱的基础上自适应确定权值, 最高融合分数所对应的标签即为该图像的识别结果. 实验结果表明, 该方法能较好地提取图像的仿射不变特征, 并且该方法在Alphanumeric, MPEG-7, GTSRB和MNIST数据库的识别准确率分别为92.2%, 96%, 90%和87.3%.

关键词: 自适应Harris角点检测; 去冗余的直线检测; 几何特征; 仿生启发特征; 仿射目标识别

基金项目: 国家自然科学基金(No.61976224, No.61976088)

中图分类号: TP391.4; TN911.73

文献标识码: A

文章编号: 0372-2112(2023)06-1607-12

电子学报URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20210450

Geometry and Bionic Fusion Feature Extraction Method for Affine Target Recognition

YU Ling-li¹, YI Qian¹, JIN Ming-yue¹, ZHOU Kai-jun²

(1. School of Automation, Central South University, Changsha, Hunan 410083, China;

2. School of Intelligent Engineering and Intelligent Manufacturing, Hunan University of Technology and Business, Changsha, Hunan 410205, China)

Abstract: In view of the problems that the target image is elongated or compressed in the horizontal or vertical direction due to different shooting angles, and the target image cannot be correctly recognized, this paper designs a feature extraction method that combines geometry and bionic vision. Aiming at geometric feature extraction, this paper improves the traditional corner and line detection algorithms, and proposes an adaptive Harris corner detection algorithm based on iterative threshold that draws on the similarity of regional pixels and a de-redundant line detection algorithm based on Hough transform. The number of corner points, the number of straight lines, and the area ratio feature vectors are taken as geometric features. Then, the bio-inspired transformation algorithm is used to extract the bionic visual features of the target image. The algorithm includes two stages. Each stage needs to perform directional edge detection and local spatial frequency detection. Calculate the Pearson correlation distance between the geometric and bionic inspired features of the target image and the features in the standard database to obtain the matching scores. The weights are adaptively determined on the basis of considering the distinguishing strength of the two characteristics of different databases. And the label corresponding to the highest fusion score is the recognition result of the image. Experimental results show that the fusion method can well extract the affine invariance features of affine images. The recognition accuracy of this method in Alphanumeric, MPEG-7, GTSRB and MNIST databases are 92.2%, 96%, 90% and 87.3% respectively.

Key words: adaptive Harris corner detection; de-redundant line detection; geometric features; bionic inspired features; affine target recognition

Foundation Item(s): National Natural Science Foundation of China (No.61976224, No.61976088)

1 引言

目标识别作为模式识别与图像处理的主要环节,在军事、工业、医学等诸多领域都起着至关重要的作用^[1,2].然而,在图像采集过程中,图像中的目标往往会由于姿态、尺度的变化而发生仿射变换^[3].因此仿射不变特征的提取在目标识别中起着重要的作用.目前的仿射不变特征存在快速但不稳定,健壮但缓慢的缺点.如HOG^[4]和2DPCA^[5]能快速处理,但其仿射不变性差,当将其用于仿射目标识别时,正确识别率低.而其他特征(如ASIFT^[6],GMbSOF^[7])虽提高了对各种几何变换的鲁棒性,但计算复杂度高.基于此,大量研究者将研究方向转至基于深度学习的目标识别^[8,9],但是深度学习在小型数据库上不够可靠且可解释性较差,导致其面向实际性较差.本文主要研究面向仿射目标识别的特征提取,通过融合几何与仿生启发特征来提升仿射目标识别性能.

几何仿射不变特征在目标识别能力、计算复杂性和稳健性上都具有较好的表现^[10],故常用几何理论来构造不变量.Kazmi等人^[11]通过实验证明了,在仿射变换下,用Harris检测器提取的角点具有几何稳定,且这些点对光照变化也很稳定.同时,研究者们发现经仿射变换后,直线、射线和直线段的像依旧分别为直线、射线和直线段^[12].且与点记号相比,使用线作为匹配特征能大量减轻特征匹配的负担^[13].为充分利用图像信息,提高仿射目标识别的准确率,高峰等人^[14]通过依次连接质心和拓展质心得到四边形,再从得到的四边形获取新的质心和拓展质心,如此迭代,构建仿射不变面积比特征向量.实验表明,面积比不变特征向量稳健性好,计算速度快,分类精度高.几何特征对扭曲变形的图像具有良好的不变性,但是其极易受到离散化错误和噪声的影响,对缩放的目标其不变性表现也较差.

由于人类初级视觉皮层能在忽视旋转、位置和缩放的情况下区别典型的线条特征,因此用于提取仿射不变特征的生物启发变换正在兴起^[15].有研究者发现使用昆虫视觉系统的四阶段生物启发视觉(Biolgy-Inspired Vision, BIV)模型^[16],能够很好的克服由于目标具有较小的物理尺寸和最小的热特征所带来的挑战.Zhou等人^[17]观察到初级视觉皮层(V1)对特定方向的光条有更强的反应,并且较高水平的神经元对图像变化具有较高的鲁棒性.虽然仿生启发特征对旋转、缩放具有较高的鲁棒性,但是其在扭曲变形上的不变性表现稍弱.

由于单项特征存在一定程度的局限性,故学者们开始通过融合多个特征来提高算法的准确度^[18].Devi等人^[19]提出一种特征级和分数级融合的方法用于生物识别以提高识别性能.除此之外,还有研究者将手背几

何特征和局部静脉模式混合特征用于人类识别^[20].由此可见,融合特征可避免单模态特征的局限性,增加识别的可靠性.

基于上述分析,本文提出一种面向仿射目标识别的几何与仿生融合特征提取方法,用于提高仿射目标识别准确度.本文主要有以下3点贡献.(1)对Harris角点检测方法进行改进,提出基于区域像素相似度的自适应Harris角点检测方法,在保证精准定位角点的同时,减少伪角点,并减少计算时间.(2)对霍夫直线检测方法进行改进,提出基于霍夫变换的去冗余直线检测方法.该方法通过分类聚集将图像中相连的像素点聚集起来,存放于数组中,用以过滤掉孤立或相连数目少的像素点.按类别依次将图像空间中的点映射到参数空间中对应的累加器,以便减少直线的重复检测,提高检测效率.(3)将几何精特征映射与仿生启发粗特征映射相结合.在考虑不同数据库两种特征区分性强弱的基础上,自适应确定权值,并用于仿射目标识别.

2 面向仿射目标识别的几何与仿生融合特征提取方法

面向仿射目标识别的几何与仿生融合特征提取方法(Geometry and Bionic Fusion Feature Extraction Method for Affine Target Recognition, AFGB)主要包括特征提取和自适应加权融合两部分,总体框架如图1所示.输入的仿射目标图像经过几何特征提取和仿生启发特征(Bionic Inspired Transform, BIT)提取两部分,得到各自的特征向量之后,与标准数据库中的特征向量进行匹配得到匹配分数,最后在评估数据库两种特征表现力强弱的基础上,自动调整加权系数,得到最终匹配得分.

在几何特征提取部分,提取仿射目标图像的三种几何特征:(1)提出基于区域像素相似度的自适应Harris角点检测方法,并提取图像的角点数 h_1 ;(2)提出基于霍夫变换的去冗余直线检测方法,并提取图像的直线数 h_2 ;(3)利用拓展质心构建四边形,提取图像的面积比向量 $[ar_1, ar_2, ar_3, ar_4]$.采用归一化处理将三种特征转换为同一量纲,构建 1×6 的几何特征向量GF(Geometry Feature), $GF=[h_1, h_2, ar_1, ar_2, ar_3, ar_4]$.

在BIT特征提取部分,对输入的目标图像进行两个阶段处理,每个阶段都包含方向边缘检测(Orientation Edge Detection, OED)^[21]和局部空间频率检测(Local Spatial Frequency Detection, LSF)^[21]两部分.其中方向边缘检测用于突出边缘信息,局部空间频率检测用于显示方向和边缘间隔.将得到的 64×64 特征图转换为 1×4096 的BIT特征向量.

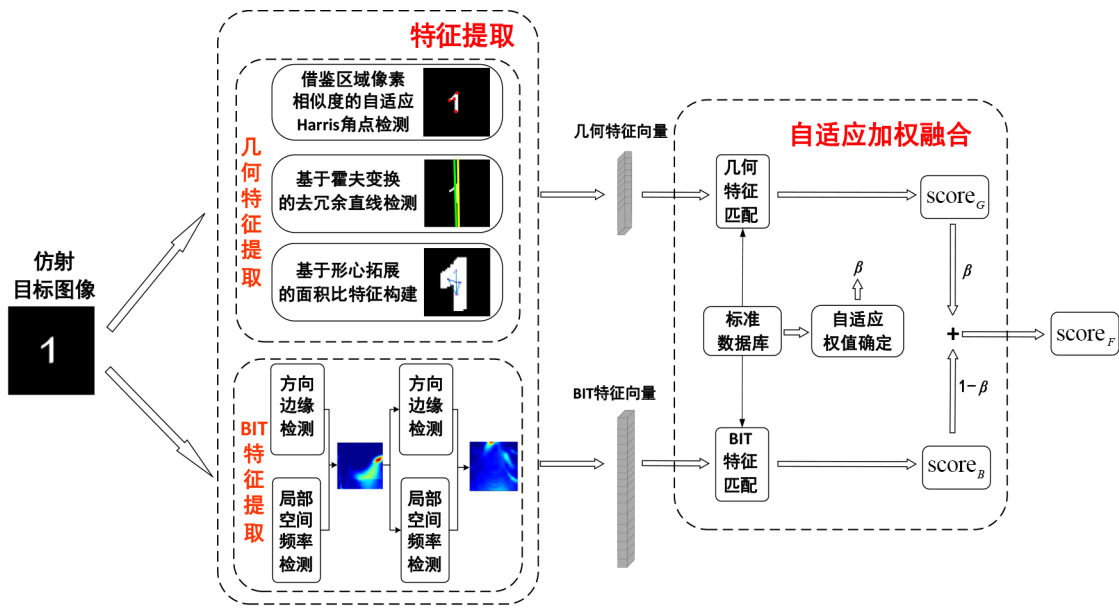


图1 面向仿射目标识别的几何与仿生融合特征提取方法总体框架

3 角点、直线、面积比向量的提取方法

3.1 借鉴区域像素相似度的自适应Harris角点检测方法

Harris角点检测方法虽然能避免相机姿态和光照的影响^[11],但是不具有尺度不变性且伪角点较多. 并且人工确定阈值的方法存在迁移性差、应用价值低的问题. 借鉴区域像素相似度的概念,提出一种改进的Harris角点检测方法,并且引入自适应阈值确定法^[22]. 具体方法如下所示.

首先求取输入图像 $I(x,y)$ 在 x 方向和 y 方向的一阶导数 I_x 和 I_y ,如式(1)所示. 之后利用高斯窗口对图像进行高斯平滑,提高图像抗噪能力,如式(2)所示. M 为Harris的相关矩阵^[11],如式(3)所示. 计算每个像素点的 R 值,如式(4)所示,并构建 R 矩阵.

$$I_x = \frac{dI(x,y)}{dx}, I_y = \frac{dI(x,y)}{dy} \quad (1)$$

$$G_s = \exp\left[-\frac{1}{2}\frac{(x^2+y^2)}{\sigma^2}\right] \quad (2)$$

$$M = G_s \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} A & B \\ B & C \end{bmatrix} \quad (3)$$

$$R = \left\{ I_x^2 \times I_y^2 - (I_x I_y)^2 \right\} - k(I_x^2 + I_y^2) \quad (4)$$

其中: σ 为高斯标准差; \otimes 为卷积运算.

当邻域像素点与中心像素点的灰度差在阈值 T 内时,认为该点为中心点. 同时,相似度计数器 C_s 加1. 在遍历完 n 个邻域后,将 C_s 在 $[2,6]$ 内的点列为候选角点,并计算其响应函数值CRF(Corner Response Func-

tion),如式(5)所示.

$$CRF = \det(M) - k \cdot \text{Tr}^2(M) \quad (5)$$

其中: $\det(M) = AC - B^2$ 为矩阵 M 的行列式; $\text{Tr}(M) = A + C$ 为矩阵 M 的迹; k 的取值范围一般为0.04~0.06.

利用迭代法计算阈值:

Step1 设置迭代终止值 $K(K=1)$ 和初值 T_0, T_0 为 R 矩阵最大值和最小值的算术平均值.

Step2 将 T_0 赋给 T_1 ,并用 R 矩阵中所有大于 T_1 的元素组成 G_1 ,所有小于 T_1 的元素组成 G_2 . 求取 G_1 和 G_2 的算术平均值 μ_1 和 μ_2 , μ 为 μ_1 和 μ_2 的算术平均值.

Step3 求 μ 与 T_1 差值的绝对值,如果其小于预定值 K ,则停止迭代阈值 T_1 ;如果大于预设值,则将 μ 的值赋给 T_1 ,重复step2和step3,直到差值小于预定值 K .

Step4 此时的 T_1 则为该图片的角点阈值.

当目标图像中的某像素点 (x,y) 的响应大于其邻域内任一像素点,同时灰度值之差在阈值 T_1 内时,认为该像素点为角点.

将改进前后的角点检测结果进行比较,如图2所示. 结果表明,Harris角点检测算法能检测到正确的角点,但是也存在大量伪角点;基于区域像素相似度的自适应Harris角点检测算法在精准定位角点位置的同时,还能有效去除伪角点.



图2 Harris角点检测改进前后的结果比较

改进后的 Harris 算法是在对像素点做了初始选择后才计算的 CRF, 该操作在一开始便去除了大量的非角点像素点, 减少了由于 CRF 的计算而导致的乘法运行次数. 时间缩短率可达 60% 以上, 如表 1 所示.

表 1 Harris 角点检测改进前后时间复杂度对比

	改进前/ms	改进后/ms	时间缩短率/%
图片 1	0.274 80	0.085 87	68.75
图片 2	0.317 61	0.054 61	82.80
图片 3	1.263 98	0.115 55	90.86

3.2 基于霍夫变换的去冗余直线检测方法

霍夫变换抗干扰能力强, 且对噪声不敏感^[23], 但是其会重复检测直线, 从而导致出现直线数量计算出错且时间复杂度增加的情况. 针对此弊端, 提出基于霍夫变换的去冗余直线检测算法. 具体方法如下所示.

首先, 利用 sobel 算子对图像进行边缘检测, 在图像空间中将各像素点按照邻接关系进行分类聚集^[24]. 通过分类聚集将图像中相连的像素点聚集在一起, 并且存放在数组中, 从而过滤掉孤立或相连数目少的像素点. 接着按类别依次将图像空间中的点映射到参数空间中对应的累加器, 如式(6)所示. 在参数空间中进行累加投票和局部峰值检测, 最后确定直线参数并映射到图像空间, 在空间中展现直线检测结果^[21].

$$\rho = x \cos \theta + y \sin \theta \quad (6)$$

其中: θ 为直线与极轴夹角, 在此处设置 θ 为 $[-89^\circ, 90^\circ]$, 步长为 0.5.

利用 MPEG-7 数据库中的正方形图片进行改进后的直线检测, 原图像如图 3(a) 所示, 检测结果如图 3(b) 所示. 按照分类聚集的方式, 该图像的边缘检测图被分为四类.

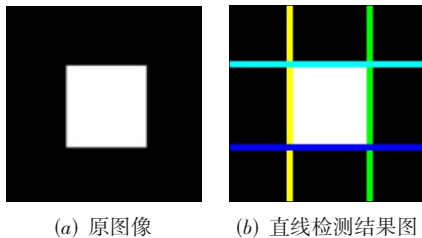


图 3 原图像及直线检测结果

改进前后的霍夫变换算法检测结果分别如图 4(a)~(d) 所示. 由此可见, 改进后的算法在精准定位目标图像每条直线的同时, 能消除冗余直线.

3.3 基于形心拓展的面积比向量构建方法

图像的形心具有仿射变换前后相对位置关系不变的性质. 然而一个图像只具有一个形心且一个形心无法构成多边形区域, 因此需要进行形心拓展^[14]. 假定

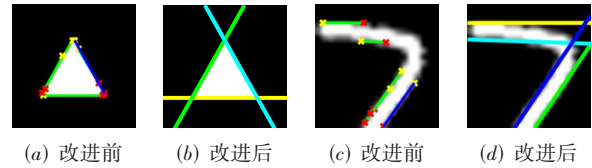


图 4 改进前后的直线检测结果比较

$c_0(x_0, y_0)$ 为输入图像 I_0 的形心, 拓展形心的求解步骤为: (1) 求取 I_0 的最大内接椭圆 e_0 ; (2) 根据区域描述子得到 e_0 的长轴半径 r_l 及长轴与 X 轴夹角 θ_l ; (3) 将 r_l 的中点坐标 $c_1(x_1, y_1)$ 定义为拓展形心, 如式(7)所示.

$$x_1 = \frac{2x_0 + r_l \cos \theta_l}{2}, y_1 = \frac{2y_0 + r_l \sin \theta_l}{2} \quad (7)$$

构造面积比特征向量的具体步骤如下:

Step1 利用形心 c_0 与拓展形心 c_1 确定的直线将目标区域划分为两个部分, 计算两部分的形心并标记为 c_2 与 c_3 ^[14], 如图 5 所示.

Step2 同理, 用直线连接 c_2 与 c_3 , 计算划分后两部分的形心并标记为 c_4 与 c_5 ; 继续上述形心计算, 提取到 12 个形心后停止迭代.

Step3 利用四边形的面积比构建向量^[14]. 根据 12 个形心, 需要计算的四边形面积为 $c_0c_2c_1c_3$, $c_0c_4c_2c_5$, $c_0c_6c_3c_7$, $c_1c_8c_2c_9$ 和 $c_1c_{10}c_3c_{11}$. 将这五个区域的面积记为 s_0, s_1, s_2, s_3, s_4 . 面积比特征向量如式(8)所示.

$$AR = [ar_1, ar_2, ar_3, ar_4], ar_i = \frac{s_i}{s_{i+1}} \quad (8)$$

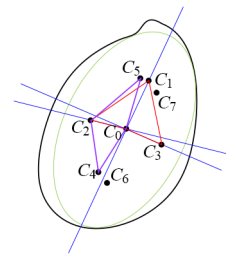


图 5 拓展形心和四边形区域构建

4 仿生启发变换(BIT)特征提取方法

4.1 方向边缘检测

方向边缘检测滤波器 $E(\theta, I)$ 由底层滤波器 Gabor 滤波器 $G_{\theta, \sigma}(x, y)$ 和水平与垂直方向的双极滤波器 $F(\theta)$ 组成^[22]. $E(\theta, I)$ 如式(9)所示.

$$\begin{aligned} E &= I_0 \otimes G_{\theta, \sigma}(x, y) \otimes F(\theta) \\ &= I_0 \otimes G_{\theta, \sigma}(x, y) \otimes [F_h(\theta) \otimes F_v(\theta)] \end{aligned} \quad (9)$$

其中: 方向角 $\theta \in [0, 180^\circ]$; 第一阶段间隔 $I \in [12.8, 89.6]$, 第二阶段间隔 $I \in [19.2, 108.8]$; $F_h(\theta)$ 和 $F_v(\theta)$ 分别为水平、垂直子滤波器.

利用 Gabor 方程对图像的边缘敏感, 但对光照不敏

感的性质^[22]来降低光照对图像的影响. 二维 Gabor 滤波器由正弦波与高斯函数相乘定义, 如式(10)所示.

$$G_{\theta,\sigma}(x,y) = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \cos\left(i\left(2\pi\frac{x'}{\lambda} + \phi\right)\right) \quad (10)$$

其中: λ 为波长; γ 为椭圆度的空间长度比; $x' = x \cos \theta + y \sin \theta$, $y' = -x \sin \theta + y \cos \theta$.

通过在不同方向上构建双极滤波器, 可检测并突出显示图像的方向边缘^[22]. 用余弦函数构建 1×3 的水平滤波器 $Hb(\theta)$, 正弦函数构建 3×1 垂直滤波器 $Vb(\theta)$. 考虑到在 $\theta = 0^\circ$, $Hb(\theta) = 0$ 时, 水平方向上的像素会被移除, 图像水平边缘会被破坏. 故将滤波器的每个元素都与设计的阶跃函数响应进行加权, 加权系数如式(11)所示.

$$S(x) = \begin{cases} 1, x > 0 \\ 0, x \leq 0 \end{cases} \quad (11)$$

根据权值因子设置阶跃函数, 水平子滤波器设置为 $Hb(\theta)$ 和 $S(Hb(\theta))$ 的点乘, 如式(12)所示. 同理, 垂直子滤波器如式(13)所示.

$$F_h(\theta) = Hb(\theta) \cdot S(Hb(\theta)) = \begin{bmatrix} -\cos(\theta) \times S(-\cos(\theta)) \\ (1 - |\cos(\theta)|) \times S(1 - |\cos(\theta)|) \\ \cos(\theta) \times S(\cos(\theta)) \end{bmatrix} \quad (12)$$

$$F_v(\theta) = Vb(\theta) \cdot S(Vb(\theta)) = \begin{bmatrix} -\sin(\theta) \times S(-\sin(\theta)) \\ (1 - |\sin(\theta)|) \times S(1 - |\sin(\theta)|) \\ \sin(\theta) \times S(\sin(\theta)) \end{bmatrix} \quad (13)$$

其中: $\sum F_v(\theta) = 1$, $\sum F_h(\theta) = 1$; $\theta \in [0, 180^\circ]$.

4.2 局部空间频率检测

当给定边缘图 E 的方向 θ 和间隔 I 时, 利用局部空间频率检测 $R(\theta, I)$ 来输出错位边缘的空间频率^[22], 如式(14)所示. 此外, 为了确保只有一条边出现时不会出现尖峰, 需将方向边缘图 E 按照 $\theta + 90^\circ$ 的方向平移间隔 I 并乘以本身. 最后对乘法图像中的所有像素求和, 求和的值按方向边缘图 E 的平方和进行归一化. 最终得到 64×64 的特征图, 将特征图转换为 1×4096 的 BIT 特征向量.

$$R(\theta, I) = \frac{\sum E \times E(\theta + 90, I)}{E} \quad (14)$$

其中: 在第一阶段, 间隔 I 的范围为输入图像大小的 $10\% \sim 70\%$; 在第二阶段, I 的范围为第一阶段特征图大小的 $15\% \sim 85\%$.

5 几何特征和仿生特征的自适应融合方法

5.1 几何特征、仿生启发特征匹配方法

为消除三种几何特征由于量纲不同造成的影响, 将角点数、直线数和面积比向量进行归一化处理, 如式(15)所示. 缩放到 $[0, 1]$ 后再构建 1×6 的几何特征向量.

$$x'_{ij} = \frac{x_{ij} - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (15)$$

其中: x_{ij} 为第 i 种几何特征的第 j 个特征值; $\min(\cdot)$ 和 $\max(\cdot)$ 为取最小值和最大值.

设 F_G 和 F_B 分别为测试图像的几何和 BIT 特征向量, T_G^k 和 T_B^k 分别为训练图像的几何和 BIT 特征向量. 采用 Pearson 相关性来测量特征向量之间的相似性, 如式(16)所示. 其中最大距离 1 表示不同图像, 反之亦然.

$$d = \text{dis}(F, T^k) = 1 - \frac{\langle (F - \bar{F}) \cdot (T^k - \bar{T}^k) \rangle}{\|F - \bar{F}\| \times \|T^k - \bar{T}^k\|} \quad (16)$$

其中: $\langle \cdot \rangle$ 表示内积.

使用匹配距离作为匹配得分, 几何特征和 BIT 特征的匹配得分分别用 score_G 和 score_B 表示, 如式(17)所示.

$$\text{score}_G = \begin{bmatrix} d_{11}^G & d_{12}^G & \cdots & d_{1j}^G \\ d_{21}^G & d_{22}^G & \cdots & d_{2j}^G \\ \vdots & \vdots & \ddots & \vdots \\ d_{i1}^G & d_{i2}^G & \cdots & d_{ij}^G \end{bmatrix}, \quad (17)$$

$$\text{score}_B = \begin{bmatrix} d_{11}^B & d_{12}^B & \cdots & d_{1j}^B \\ d_{21}^B & d_{22}^B & \cdots & d_{2j}^B \\ \vdots & \vdots & \ddots & \vdots \\ d_{i1}^B & d_{i2}^B & \cdots & d_{ij}^B \end{bmatrix}$$

5.2 自适应权值确定方法

采用经典的加权模型对匹配分数进行融合, 得到融合得分 score_F , 如式(18)所示.

$$\text{score}_F = \beta \times \text{score}_G + (1 - \beta) \times \text{score}_B = \begin{bmatrix} d_{11}^F & d_{12}^F & \cdots & d_{1j}^F \\ d_{21}^F & d_{22}^F & \cdots & d_{2j}^F \\ \vdots & \vdots & \ddots & \vdots \\ d_{i1}^F & d_{i2}^F & \cdots & d_{ij}^F \end{bmatrix} \quad (18)$$

其中: β 为加权系数, $\beta \in [0, 1]$.

不同数据库的几何特征和仿生启发特征的图片区分性存在相对优劣的情况, 例如 Alphanumeric 数据库各个图片的几何特征比较相似, 导致其几何特征在各个图片中的区分性差. 而其仿生启发特征的表现力较好.

考虑到上述分析, 提出一种自适应的加权融合. 计算训练集图像几何特征的平均距离 d_1 和仿生特征平均距离 d_2 , 分别如式(19)和式(20)所示. 距离 d_1 越大, 则

表示该数据库的几何特征区分性越好,故给其权重 β 越大,如式(21)所示.

$$d_1 = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=i+1}^n \text{dis}(T_G^i, T_G^j) \quad (19)$$

$$d_2 = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=i+1}^n \text{dis}(T_B^i, T_B^j) \quad (20)$$

$$\beta = \frac{d_1}{d_1 + d_2} \quad (21)$$

最后,通过最近邻算法分类到对应最小距离的参考图像的第 k 类,如式(22)所示.

$$O_F = \underset{k \in K}{\text{argmin}} \{ \text{score}_F \} \quad (22)$$

6 实验结果与分析

6.1 数据库和实验环境

实验图像来自 Alphanumeric(字母数学)、MPEG-7、GTSRB(交通标志牌)和 MNIST(手写数字)数据库,部分样式如图6所示.

6.2 几何特征的选择性与不变性分析

6.2.1 角点和霍夫直线检测实验

为验证角点数和直线数存在仿射选择性与不变性,现对原图像进行不同尺度下的仿射变换,并将其检测结果和原图像的结果进行比较,结果如图7、表2、表3所示.从表2和表3的粗体差值能看出,同一类目标图像的角点数与霍夫直线数的差在 ± 3 的范围内波动.且除手写字体这种边缘不明显的图像外,其他边缘平稳图像的角点数与霍夫直线数的差更是稳定在 ± 1 范围内.而不同类的目标图像,其角点数或霍夫直线数的差在 ± 4 范围外.该实验充分证明了用角点数与霍夫直线数作为仿射目标的特征并用于识别是可靠的.



(a) Alphanumeric 数据库



(b) MPEG-7 数据库



(c) GTSRB 数据库

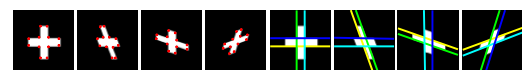


(d) MNIST 数据库

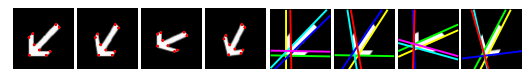
图6 实验数据库



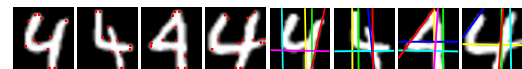
(a) 数字1



(b) 十字标志



(c) “靠右行驶”标志



(d) 手写数字4

图7 仿射目标的角点检测和霍夫直线检测结果

表2 原图像和仿射图像的角点数量及差值表

图片名称	原图像	仿射图像1		仿射图像2		仿射图像3		仿射图像4	
	Harris	Harris	差值	Harris	差值	Harris	差值	Harris	差值
数字1	4	4	0	4	0	4	0	4	0
十字标志	12	12	0	12	0	12	0	12	0
“靠右行驶”	9	9	0	9	0	9	0	9	0
手写数字4	12	11	-1	12	0	15	+3	14	+2

表3 原图像和仿射图像的直线数量及差值表

图片名称	原图像	仿射图像1		仿射图像2		仿射图像3		仿射图像4	
	Hough	Hough	差值	Hough	差值	Hough	差值	Hough	差值
数字1	2	2	0	2	0	2	0	2	0
十字标志	4	4	0	4	0	4	0	4	0
“靠右行驶”	6	5	-1	6	0	5	-1	6	0
手写数字4	5	5	0	6	+1	5	0	5	0

6.2.2 面积比实验

为直观地表明面积比特征向量具有良好的仿射不变性与选择性,现将各图像与其仿射变换图像的面积比向量绘制成特征曲线,如图8所示.分别观察四幅图的曲线一致性,除了在手写数字4处出现了一处比较明显的差别外,整体的变化趋势协同性高.联合观察四幅图,可以看出,不同类图像的趋向一致性存在比较大的不同.上述分析表明面积比特征向量对仿射变换有较强的鲁棒性,而且在不同类图像之间有较强的区分性.因此,可利用面积比特征向量去构建目标图像的几何特征用于仿射识别.

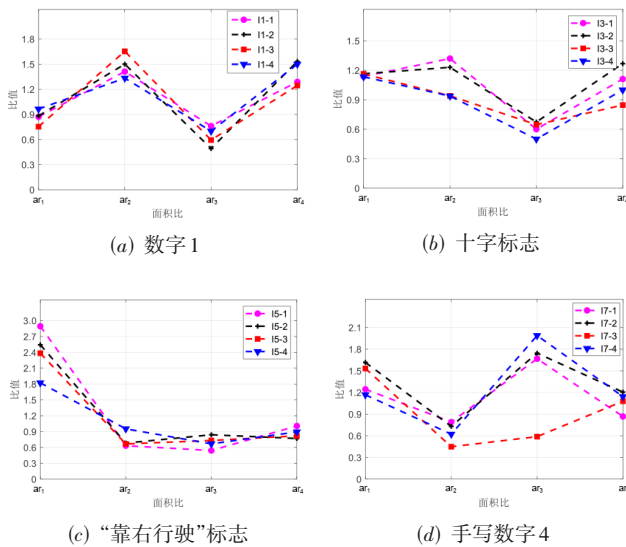


图8 仿射目标的面积比向量结果

6.2.3 采用MDS对几何特征进行可视化分析

6.2.1和6.2.2节从主观方面论证了几何特征的不变性与选择性.本节用非经典多维尺度变换(non-classical MultiDimensional Scaling, MDS)从客观方面进行证明.MDS是通过度量欧氏距离,来保留数据点对间相似性(或不相似性)的降维方法^[14].

图9中每组图片的左边是每个数据库的Shepard图,该图显示了非计量缩放的距离是如何近似于差距的.在该图中, X 轴对应于输入接近度, Y 轴对应于MDS距离和转换后的输入接近度.为了更准确地重构距离,图9中每组图片的右边给出了三维特征数据分布图.该图中,蓝色的点表示几何不变特征数据,红色的数字表示目标图像的类别.对于同类的仿射目标,由于其相似度高,故其距离应该相对较近.

图9(a)所示的Alphanumeric数据库的图像分类较为聚集,即部分数字字母图像的几何特征类似,区分度较低;图9(c)所示的GTSRB数据库的图像分类较为清晰,即交通标志牌的几何特征具有较强的选择性.整体

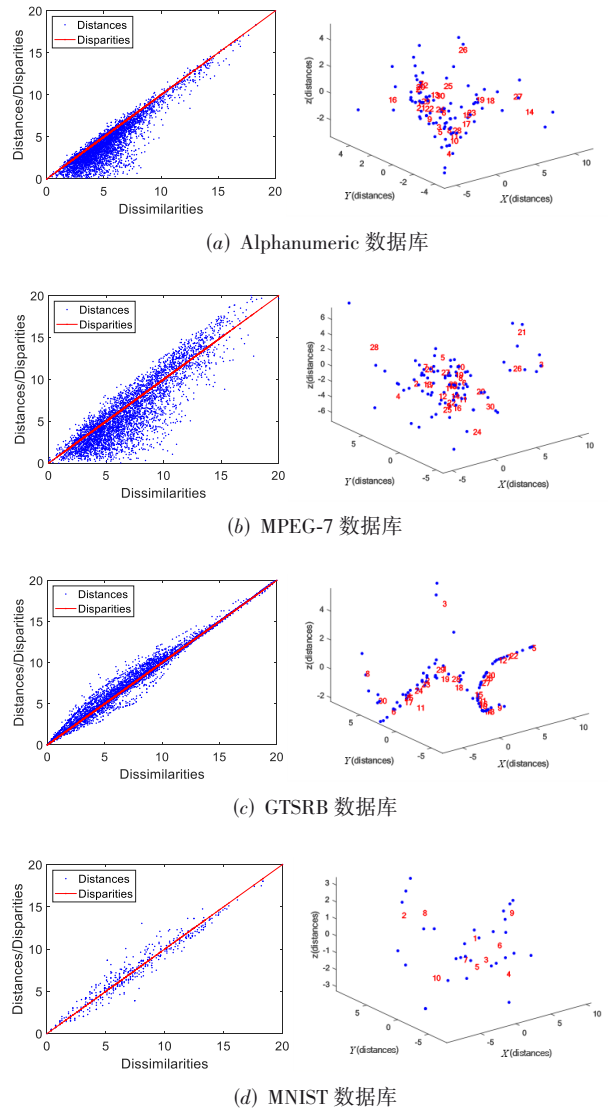


图9 仿射目标几何特征的MDS统计实验结果

来说,几何不变特征具有一定选择性.

6.3 BIT特征的选择性与不变性分析

从各数据库中各选一幅图像进行不同剪切因子下的仿射变换,并将其BIT特征图展示在图10中.

为保证比较的公平性,同样采取MDS评价BIT特征的选择性与不变性,结果如图11所示.对比图9和图11知:Alphanumeric数据库和MPEG-7数据库的BIT特征的表现优于其几何特征;而GTSRB数据库的几何特征则展现了更好的仿射不变性与选择性;MNIST数据库的几何特征与BIT特征的优劣程度则较为相似.

6.4 分类方法比较

为提高识别的准确率并降低识别的时间复杂度,现将各分类方法分别用于该识别方法.图12展示了采用不同分类方法的识别准确率对比.分类方法包括最近邻(K-Nearest Neighbor, KNN)、支持向量机(Support

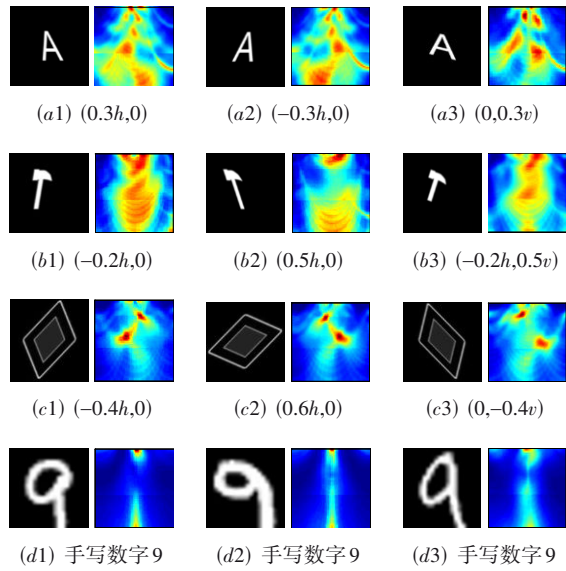


图 10 仿射目标在不同剪切因子下的 BIT 特征图

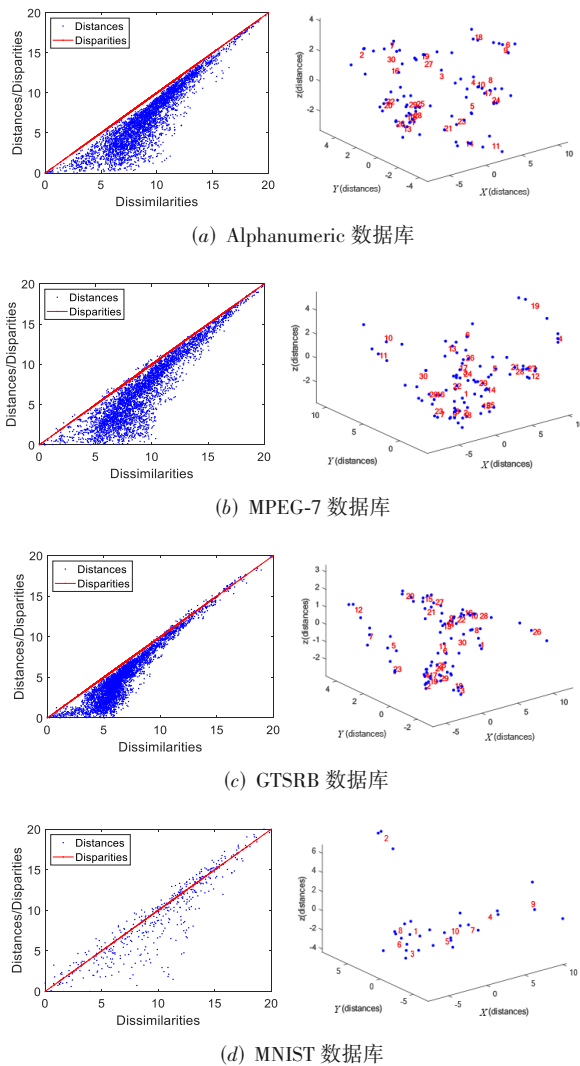


图 11 仿射目标 BIT 特征的 MDS 实验结果

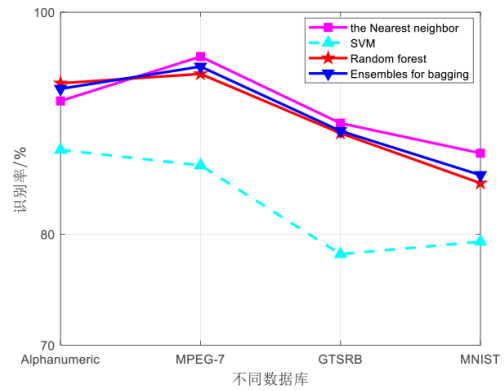


图 12 各分类方法的识别准确率比较

Vector Machine, SVM)、随机森林(Random Forest, RF)和装袋方法(Ensemble for Bagging, EFB). 从图看出,除在 Alphanumeric 数据库中 FB 和 RFE 的识别准确率相比于 KNN 略胜一筹外,在其他数据库中, KNN 在识别准确度方面的优势是显而易见的. 并且从 KNN 的性质来看, KNN 的模型易理解, 通常不需要过多调节就能得到不错的性能, 故构建 KNN 的速度通常很快. 因此从整体性能考虑, KNN 是最适合该算法的分类方法.

6.5 融合实验分析

为证明几何和仿生融合特征具有更高的仿射不变性, 从 MPEG-7 数据库中选 6 类目标图像进行实验, 对每一类目标图像随机进行三个范围为 $(-0.5, 0.5)$ 的剪切因子计算.

表 4 中的每一组数字分别代表仿射变换图像与原始图像匹配的相关系数与欧氏距离, 相关系数大和欧式距离小均代表图像直接相似度高. 从表 4 的粗体值能看出同类别图像的融合相关系数接近于 1, 欧氏距离接近于 0. 而不同类别的仿射变换图像, 相关系数较低且欧氏距离较大. 整体来看, 该融合方法对于同类和不同类的误判率低, 能精准地识别仿射目标图像.

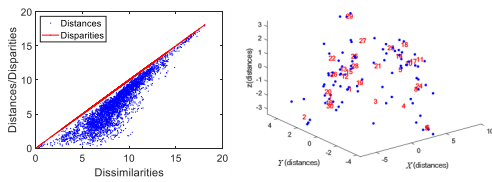
同样使用 MDS 分析各数据库融合特征的不变性与选择性, 实验结果如图 13 所示. 将图 9、图 11、图 13 进行比较, 可见在采用融合特征下, 同类的仿射目标更接近, 不同类的仿射目标距离较远. 这表明, 比起采用单一特征进行识别, 融合的方法保留了更多的信息来区分不同类别的目标图像.

6.6 本文所提方法与其他仿射目标识别方法的准确率和计算时间比较

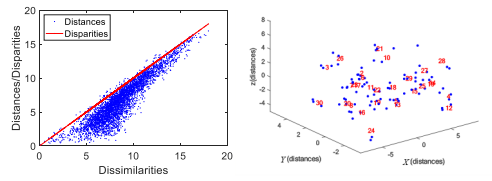
为了证明本文所提方法比其他方法的识别率更高且不需要大量的训练样本, 本节将该方法与其他一些经典的仿射目标识别方法进行对比. 从 Alphanumeric, MPEG-7 和 GTSRB 数据库中各随机选取 30 幅未进行仿射变换的图像作为标准集; 将对每幅标准图像进行 $(0, 0)$, $(-0.3h, 0)$, $(0, -0.3v)$, $(0.2h, 0.3v)$ 和 $(-0.2h,$

表 4 MPEG-7 数据库分数的部分实验结果

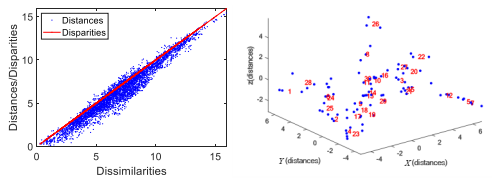
	Bell	Bat	Apple	Pencil	Tree	Bird		Bell	Bat	Apple	Pencil	Tree	Bird
Bell-1	0.993 0	0.713 6	0.622 3	0.376 5	0.755 2	0.769 2	Pencil-1	0.385 1	0.372 2	0.369 9	0.995 8	0.453 0	0.417 5
	0.037 1	0.168 1	0.156 3	0.223 0	0.118 8	0.106 2		0.302 6	0.479 4	0.136 2	0.007 9	0.418 7	0.660 1
Bell-2	0.841 5	0.677 5	0.642 8	0.411 0	0.814 4	0.823 7	Pencil-2	0.396 9	0.386 5	0.507 9	0.988 8	0.542 6	0.454 6
	0.085 3	0.211 5	0.132 4	0.206 3	0.136 3	0.102 1		0.289 6	0.509 1	0.126 9	0.009 3	0.470 1	0.438 0
Bell-3	0.987 2	0.733 1	0.705 9	0.523 8	0.721 7	0.781 5	Pencil-3	0.286 0	0.286 0	0.300 5	0.890 0	0.331 6	0.143 4
	0.065 1	0.257 3	0.106 0	0.125 5	0.165 3	0.136 4		0.350 5	0.540 2	0.141 2	0.016 0	0.568 7	0.495 8
Bat-1	0.742 0	0.982 9	0.588 8	0.423 8	0.784 9	0.850 5	Tree-1	0.815 4	0.782 2	0.759 3	0.459 0	0.923 8	0.764 6
	0.116 0	0.029 4	0.189 4	0.198 3	0.109 2	0.083 8		0.100 4	0.102 4	0.140 5	0.238 2	0.089 0	0.138 2
Bat-2	0.670 5	0.907 3	0.550 0	0.377 7	0.756 9	0.789 1	Tree-2	0.759 2	0.720 5	0.676 0	0.429 6	0.995 6	0.851 1
	0.139 7	0.067 7	0.212 7	0.203 0	0.123 9	0.108 5		0.102 7	0.124 4	0.158 2	0.216 0	0.002 3	0.154 3
Bat-3	0.723 3	0.962 6	0.533 0	0.392 7	0.811 3	0.657 9	Tree-3	0.785 3	0.779 3	0.733 7	0.482 8	0.919 7	0.772 5
	0.173 0	0.047 9	0.269 6	0.246 5	0.123 1	0.114 1		0.104 5	0.066 0	0.116 4	0.186 3	0.085 1	0.136 0
Apple-1	0.654 6	0.584 9	0.994 6	0.458 9	0.646 5	0.663 6	Bird-1	0.811 6	0.865 2	0.761 2	0.405 7	0.783 8	0.960 0
	0.162 4	0.392 0	0.001 2	0.102 1	0.255 9	0.407 9		0.090 9	0.107 0	0.145 4	0.243 5	0.102 8	0.064 9
Apple-2	0.648 0	0.495 7	0.907 9	0.386 0	0.718 9	0.722 0	Bird-2	0.812 3	0.780 4	0.655 9	0.377 1	0.806 5	0.994 4
	0.187 6	0.525 7	0.031 5	0.137 7	0.392 7	0.324 3		0.114 1	0.160 1	0.220 9	0.293 7	0.147 6	0.070 0
Apple-3	0.638 4	0.426 4	0.954 5	0.433 2	0.706 0	0.703 3	Bird-3	0.722 2	0.834 6	0.706 8	0.463 0	0.772 9	0.900 0
	0.193 9	0.517 8	0.015 9	0.136 3	0.246 5	0.451 9		0.145 0	0.201 6	0.215 7	0.321 3	0.174 2	0.086 5



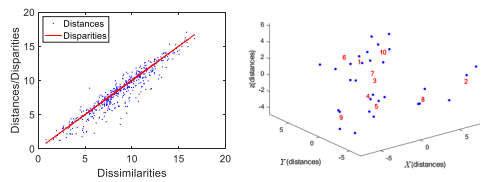
(a) Alphanumeric 数据库



(b) MPEG-7 数据库



(c) GTSRB 数据库



(d) MNIST 数据库

图 13 仿射目标融合特征的 MDS 实验结果

-0.3v) 仿射变换后的 5 幅图像作为训练集; 将随机进行 (-0.7h, +0.7v) 范围内仿射变换的 15 幅图像作为测试集. 在 MNIST 数据库选每类图像的前 5 幅为训练集, 后 15 幅为测试集. 故 Alphanumeric, MPEG-7 和 GTSRB 数据库各有 150 幅训练图像和 450 幅测试图像, MNIST 数据库有 50 幅训练图像和 150 幅测试图像.

其他算法包括: 生物启发变换算法 (Biological Inspired Transformation, BIT)、几何特征提取 (Geometric Feature, GF)、方向梯度直方图 (Histogram of Oriented Gradient, HOG)、二维主成分分析方法 (2D Principal Components Analysis, 2DPCA)、尺度不变特征变换算法 (Scale-Invariant Feature Transform, SIFT) 和加速稳健特征算法 (Speeded Up Robust Features, SURF). 比较结果如表 5 和图 14 所示. 融合算法在 MPEG-7, Alphanu-

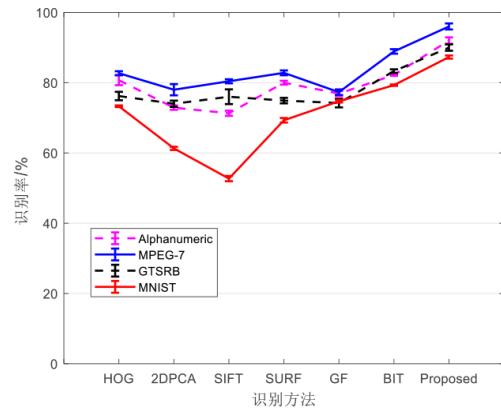


图 14 各仿射目标识别方法的误差棒曲线

meric, GTSRB 和 MNIST 数据库的平均正确识别率为 92.2%, 96%, 90% 和 87.3%, 高于 HOG, 2DPCA 等其他方法. 这表明该方法能够弥补 HOG, 2DPCA 等仿射不变特征的低识别率, 为仿射目标图像的识别提供新的参考.

在计算复杂度方面, 对于 450 幅测试图像的 Alphanumeric, MPEG-7 和 GTSRB 数据库, 融合算法的识别时间分别为 6 107.2 s, 6 601.3 s 和 6 041.8 s; 对于 150 幅训练图像的 MNIST 数据库, 识别时间为 1 955.9 s. 由于仿生启发特征的提取时间较长, 融合算法的时间比 SIFT

以外的方法都要长. 因此, 下一步的工作需从减少 BIT 特征的提取阶段、简化 BIT 特征的提取步骤等方面着手.

最后, 分析在不同剪切因子下, 不同识别方法的识别准确率变化情况. 对于 Alphanumeric 和 MPEG-7 数据库, 图 15(a1)和(b1)、图 15(a2)和(b2)、图 15(a3)和(b3)分别展示了水平、垂直、水平和垂直剪切因子在 (-0.7, +0.7) 范围时识别准确率的变化. 从图 15 整体来看, AFGB 方法对于仿射变换图像的识别准确率最高.

表 5 不同特征提取方法在不同数据库中仿射变换图像的秩 1 识别准确率和识别计算时间

数据集	秩 1 识别准确率/%				计算时间/s			
	Alphanumeric	MPEG-7	GTSRB	MNIST	Alphanumeric	MPEG-7	GTSRB	MNIST
HOG	82.7	80.7	76.2	73.3	3.35	3.38	3.79	1.34
2DPCA	72.9	78.0	74.0	61.3	29.8	34.7	34.8	13.7
SIFT	71.3	80.4	76.0	52.7	19 532.7	19347.0	20 766.3	8 754.4
SURF	80.0	82.8	74.9	69.3	855.8	834.4	846.9	86.2
GF	76.9	77.3	74.2	74.7	143.3	144.2	153.3	50.4
BIT	82.2	88.9	83.3	79.3	5 962.7	6 465.1	5 864.2	1 938.2
本文	92.2	96.0	90.0	87.3	6 107.2	6 601.3	6 041.8	1 955.9

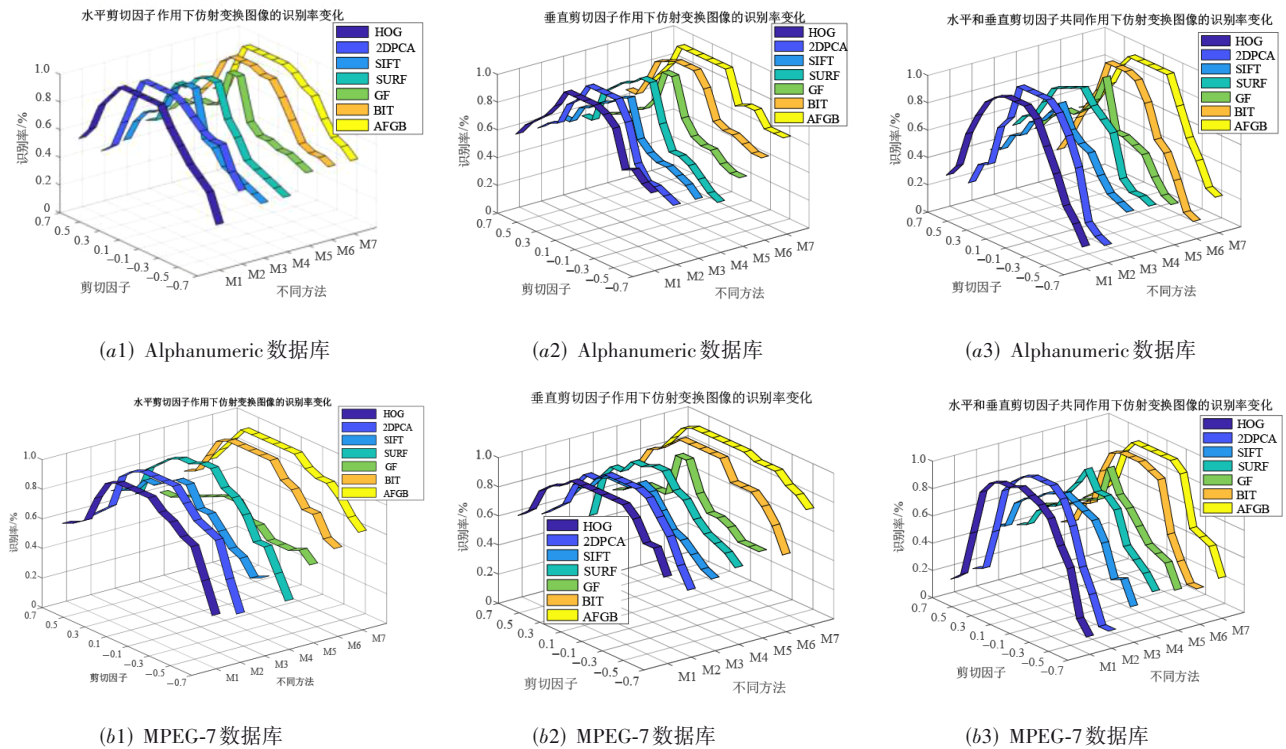


图 15 不同仿射因子下各仿射目标识别方法的识别准确率变化

7 结论

针对仿射目标图像识别准确率较低、识别时间过长、训练样本需求过多的问题,本文提出了一种自适应加权融合的方法.通过融合几何匹配得分与仿生启发匹配得分,将几何精特征映射与仿生启发粗特征映射相结合.该方法在充分利用几何特征的对扭曲变形的图像具有一定程度不变性的同时,结合仿生启发特征来规避其弱尺度不变性,充分发挥仿生启发特征对缩放、旋转等变换的鲁棒性.

考虑到Harris角点检测方法伪角点过多,对其数量的运用会降低识别准确率,本文在借鉴区域像素相似度的基础上,引入自适应阈值确定方法,从而改进Harris角点检测.实验结果表明,该方法能精准定位角点,有效去除伪角点,并且还能缩短角点检测时间.同样,考虑到利用霍夫变换检测直线时,冗余直线过多及直线数量特征的误判率增大的问题,本文提出一种基于霍夫变换的去冗余直线检测方法.实验结果表明,该方法在精准定位目标图像中每一条直线的同时还能消除冗余的直线,提高了直线检测的效率.

考虑到不同数据库的几何特征和仿生启发特征表现力强弱不同,本文提出根据不同数据库的自适应权重确定法.实验结果表明,将几何特征和仿生启发特征融合起来,既规避了Alphanumeric数据库中几何特征的弱选择性,也规避了GTSRB数据库中仿生启发特征的弱选择性.

总体来说,该方法在大仿射环境下,识别准确度远高于HOG和SURF等算法,能够有效地提高仿射目标图像的识别准确率,识别时间上也优于SIFT特征,并且该方法不需要大量的训练样本,避免了机器学习等方法的任务繁重性.

参考文献

- [1] WANG S J, JIANG F, ZHANG B, et al. Development of UAV-based target tracking and recognition systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(8): 3409-3422.
- [2] 翟颖,陈渤.基于稳健变分自编码模型的雷达高分辨距离像目标识别算法[J].电子学报,2020,48(6):1149-1155.
ZHAI Y, CHEN B. Robust variational auto-encoder for radar hrrp target recognition[J]. Chinese Journal of Electronics, 2020, 48(6): 1149-1155. (in Chinese)
- [3] HOSNY K M. New set of gegenbauer moment invariants for pattern recognition applications[J]. Arabian Journal for Science and Engineering, 2014, 39(10): 7097-7107.
- [4] WANG H, WEI S B, FANG B F. Facial expression recognition using iterative fusion of MO-HOG and deep features [J]. The Journal of Supercomputing, 2020, 76(5): 3211-3221.
- [5] YANG X F, WANG W S, LIU L M, et al. Robust 2DPCA by T1 criterion maximization for image recognition[J]. IEEE Access, 2021, 9: 7690-7700.
- [6] MOREL J M, YU G S. ASIFT: A new framework for fully affine invariant image comparison[J]. SIAM Journal on Imaging Sciences, 2009, 2(2): 438-469.
- [7] MAIER J, HUMENBERGER M, MURSCHITZ M, et al. Guided matching based on statistical optical flow for fast and robust correspondence analysis[C]//European Conference on Computer Vision. Virtual Conference: Springer, 2016, 9911: 101-117.
- [8] HUANG T, ZHANG Q X, LIU J B, et al. Adversarial attacks on deep-learning-based SAR image target recognition [J]. Journal of Network and Computer Applications, 2020, 162(12): 102632.
- [9] 郝彤,赵杰.面向双曲线形态的探地雷达图像识别技术综述[J].电子学报,2019,47(6):1366-1372.
HAO T, ZHAO J. A brief review of the hyperbola signature recognition techniques for ground penetrating radar[J]. Chinese Journal of Electronics, 2019, 47(6): 1366-1372. (in Chinese)
- [10] 胡欣,许新芳,陈娇.仿射不变性特征提取在目标识别中的应用[J].兵工自动化,2013(8):84-87.
HU X, XU X F, CHEN J. Affine invariability feature extraction in application of target recognition[J]. Ordnance Industry Automation, 2013(8): 84-87. (in Chinese)
- [11] KAZMI W, ANDERSEN H J. A comparison of interest point and region detectors on structured, range and texture images [J]. Journal of Visual Communication and Image Representation, 2015, 32: 156-169.
- [12] 叶正麟,安晓虹.线性代数变换与仿射几何变换[J].高等数学研究,2021,24(1):1-4.
YE Z L, AN X H. Algebra transformation and affine transformation[J]. Studies in College Mathematics, 2021, 24(1): 1-4. (in Chinese)
- [13] PATOOMMAKESORN K, VIGNAT F, VILLENEUVE F. A new straight line matching technique by integration of vision-based image processing[J]. Procedia CIRP, 2016, 41: 777-782
- [14] 高峰,文贡坚.利用仿射几何的仿射不变特征提取方法[J].中国图象图形学报,2011,16(3):389-397.
GAO F, WEN G J. Affine invariant feature extraction

based on affine geometry[J]. Chinese Journal of Image and Graphics, 2011, 16(3): 389-397. (in Chinese)

- [15] LU Y F, QIAO H, LI Y, et al. Image recommendation based on a novel biologically inspired hierarchical model [J]. Multimedia Tools and Applications, 2017, 77(4): 4323-4337.
- [16] KARIMI H S, MOHAMMADI K. Rotational invariant biologically inspired object recognition[J]. IET Image Processing, 2020, 14(15): 3762-3773.
- [17] ZHOU X, ZHOU K J, SHEN L. Rotation and translation invariant palmprint recognition with biologically inspired transform[J]. IEEE Access, 2020, 8: 80097-80119.
- [18] MOYOU M, ANAND R, JOHN C, et al. A grassmannian graph approach to affine invariant feature matching[J]. IEEE Transactions on Image Processing, 2020, 29: 3374-3387.
- [19] DEVI D R, KATTAMURI N R. Sub-band-based feature fusion and hybrid fusion approaches for multimodal biometric identification[J]. International Journal of Biometrics, 2020, 12(4): 357-376.
- [20] PITITHEERAPHAB Y, THONGPANCE N, AOYAMA H, et al. Vein pattern verification and identification based on local geometric invariants constructed from minutia points and augmented with barcoded local feature[J]. Applied Sciences, 2020, 10(9): 3192.
- [21] 余伶俐, 夏旭梅, 周开军, 等. 基于仿生视觉的图像RST不变属性特征提取方法[J]. 仪器仪表学报, 2017, 38(4): 985-995.
YU L L, XIA X M, ZHOU K J, et al. RST-invariant feature extraction method inspired by bionic visual perception[J]. Chinese Journal of Scientific Instrument, 2017, 38(4): 985-995. (in Chinese)
- [22] WANG Z C, LI R, SHAO Z H, et al. Adaptive Harris corner detection algorithm based on iterative threshold[J]. Modern Physics Letters B, 2017, 31(15): 1750181.
- [23] AGGARWAL N, KARL W C. Line detection in images through regularized Hough transform[J]. IEEE Transactions on Image Processing, 2006, 15(3): 582-591.
- [24] 宋晓宇, 袁帅, 郭寒冰, 等. 基于自适应阈值区间的广义Hough变换图形识别算法[J]. 仪器仪表学报, 2014, 35(5): 1109-1117.
SONG X Y, YUAN S, GUO H B, et al. Pattern identification algorithm with adaptive threshold interval based extended Hough transform[J]. Chinese Journal of Scientific Instrument, 2014, 35(5): 1109-1117. (in Chinese)

作者简介



余伶俐 女, 1983年生, 江西景德镇人. 中南大学自动化学院教授. 主要研究方向为移动机器人导航规划、视觉感知处理.
E-mail: llyu@csu.edu.cn



易情 女, 1999年生, 湖南长沙人. 中南大学自动化学院硕士研究生. 主要研究方向为图像处理、仿生视觉.
E-mail: 204612175@csu.edu.cn



金鸣岳(通讯作者) 女, 1996年生, 河北邯郸人. 中南大学自动化学院硕士研究生. 主要研究方向为仿生视觉感知启发下的不变特征提取方法.
E-mail: jinmingyue258@163.com



周开军 男, 1979年生, 湖南常德人. 湖南工商大学智能工程与智能制造学院教授. 主要研究方向为机器视觉、图像处理.
E-mail: zkj@hutb.edu.cn