

# 基于黄金正弦混沌斑鬣狗优化算法的高光谱波段选择

戴天虹<sup>1</sup>, 孙春雪<sup>1</sup>, 黄建平<sup>1</sup>, 谢千程<sup>1</sup>, 丛士杰<sup>1</sup>, 黄新望<sup>1</sup>, 李克新<sup>2\*</sup>

<sup>1</sup>东北林业大学机电工程学院, 黑龙江 哈尔滨 150040;

<sup>2</sup>无锡科技职业学院人工智能学院, 江苏 无锡 214000

**摘要** 波段选择是降低高光谱数据维度, 减少数据过多冗余的有效手段, 是高光谱影像像素分类的重要前提。其本质上是一个复杂的组合优化问题, 用传统的搜索方法不易得到满意的解。针对上述问题, 提出了一种结合黄金正弦和混沌斑鬣狗算法(GSSHO)的高光谱波段选择方法。首先, 使用混沌策略初始化斑鬣狗种群, 提高种群的随机性和多样性; 然后, 用黄金正弦算法改进原始斑鬣狗算法搜索个体位置更新方式, 提高算法的全局搜索能力; 最后, 设计结合分类精度和波段个数的适应度函数, 对算法优化性能进行评价。在高光谱遥感数据集上, 将该方法与其他先进优化算法进行比较, 实验结果表明, 该方法所选波段个数接近原波段的 1/10, 对于 Pavia Centre 数据集分类精度高达 99.08%, 优于其他对比方法, 能以更合理的收敛方向找到最优解, 所选波段数更少, 分类精度更高, 是一种高效的波段选择方法。

**关键词** 成像系统; 波段选择; 斑鬣狗优化算法; 黄金正弦算法; 混沌策略

中图分类号 TP751

文献标志码 A

DOI: 10.3788/LOP202259.1030003

## Hyperspectral Wave Band Selection Based on Golden Sine and Chaotic Spotted Hyena Optimizer Algorithm

Dai Tianhong<sup>1</sup>, Sun Chunxue<sup>1</sup>, Huang Jianping<sup>1</sup>, Xie Qiancheng<sup>1</sup>, Cong Shijie<sup>1</sup>,  
Huang Xinwang<sup>1</sup>, Li Kexin<sup>2\*</sup>

<sup>1</sup>College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040,  
Heilongjiang, China;

<sup>2</sup>College of Artificial Intelligence, Wuxi Vocational College of Science and Technology, Wuxi 214000, Jiangsu, China

**Abstract** Wave band selection is an effective means to reduce the dimension and much redundancy of hyperspectral data, and it is an important prerequisite for pixel classification of hyperspectral images. It is a complex combinatorial optimization problem in essence, and it is difficult to get satisfactory solution by traditional search methods. To solve the above problems, a method of hyperspectral band selection based on golden sine and chaotic the spotted hyena optimization algorithm (GSSHO) is proposed. Firstly, chaos strategy is used to initialize the spotted hyena population to improve the randomness and diversity of the population. Secondly, golden sine algorithm is used to improve the original spotted hyena optimization (SHO) algorithm to search individual position update mode to improve the global search ability of the algorithm. Finally, a fitness function combining classification accuracy and band number is

收稿日期: 2021-03-26; 修回日期: 2021-04-20; 录用日期: 2021-05-18

基金项目: 中央高校基本科研业务费专项资金(2572019CP17, 2572019CP19)、黑龙江省自然科学基金(C201414, TD2020C0001)、哈尔滨市科技创新人才项目(2014RFXXJ086)

通信作者: \*dillonlkx@163.com

designed to evaluate the optimization performance of the algorithm. On hyperspectral remote sensing data sets, this method is compared with other advanced optimization algorithms. The experimental results show that the number of bands selected in this algorithm is close to one tenth of the original band, and the classification accuracy of Pavia Centre data set is up to 99.08%, which is better than those of other comparison methods. It can find the optimal solution with more reasonable convergence direction, and the number of selected bands is less, and the classification accuracy is higher. It is an efficient method for selecting wave band.

**Key words** imaging systems; wave band selection; spotted hyena optimizer algorithm; golden sine algorithm; chaos strategy

## 1 引言

遥感技术能够提供丰富的地表信息,被广泛应用到农业监测、资源勘探、灾害预警等众多领域中。与传统遥感技术相比,高光谱遥感技术能够提供更多的光谱空间信息,有更高的光谱分辨率和连续而完整的光谱特征曲线,在目标识别和分类领域具有更大的优势<sup>[1]</sup>。然而,更高的光谱分辨率带来了更大的波段采集密度,使得相邻波段间光谱相关性增大,所获取的光谱图像更加相近,导致高光谱影像数据冗余信息增多,数据存储量变大,处理过程更加复杂<sup>[2-3]</sup>。同时,在高光谱影像分类识别中,当光谱特征超过一定数量时,可能会出现分类精度不升反降的情况,即“Hughes”现象<sup>[4]</sup>。因此,降低光谱空间的维度,剔除过多的冗余信息,实现高光谱影像波段优选是当前研究的热点。

高光谱数据的降维方法包括两类:一类是特征提取,即通过某种投影方式将光谱数据从较高维映射到较低维;另一类是波段选择,即从原始的高维光谱特征空间中筛选出若干特征波段构成波段子集。相比特征提取,波段选择没有经过任何变换,保留了图像原有特征的物理意义,更具有结果解释性,因而受到更广泛的关注和应用<sup>[5]</sup>。高光谱波段选择本质上是一种复杂的组合优化问题,基本方法是依据特定评价准则,利用某种搜索策略寻找信息量大、独立性强的波段子集<sup>[6]</sup>。元启发式优化算法有较好的目标寻优能力,在特征选择领域受到广泛的关注,常用的有遗传算法<sup>[7-8]</sup>、粒子群算法<sup>[9-10]</sup>、杜鹃搜索算法<sup>[11]</sup>等:遗传算法虽然能够解决特征数量大、遍历不全面的问题,但其收敛速度较慢;粒子群算法虽然易于实现,收敛速度快、效率高,但容易陷入局部最优,寻优效果不佳;杜鹃搜索算法虽然参数少、速度快、通用性更好,但不能很好地均衡全局性与收敛性。因此,需要不断寻求更优秀的优化算法来解决波段选择问题。

斑鬣狗优化(spotted hyena optimizer,SHO)算法是印度学者 Dhiman 与 Kumar<sup>[12]</sup>于 2017 年提出的一种新型群智能优化算法,模拟群居食肉动物斑鬣狗的群居生活和觅食习性,从而达到优化搜索的目的。该算法具有较强的自适应能力,且设置参数少、稳定性强,被用于解决复杂非线性约束工程<sup>[13]</sup>、比例积分微分(PID)优化<sup>[14]</sup>和特征选择<sup>[15]</sup>等问题,展现出较强的优化能力。本文将 SHO 算法用于高光谱波段选择研究中,虽然该算法在大多数优化问题上有着较强的寻优能力,但是在求解高维优化问题时,仍存在收敛速度慢、易陷入局部最优等问题。基于此,将黄金正弦算法与 SHO 算法相结合,利用黄金正弦算法较强的遍历能力,将其独特的位置更新方式引入到 SHO 算法中,提高 SHO 算法的收敛速度、全局搜索能力。同时,利用混沌策略初始化斑鬣狗种群,结合锦标赛策略,进一步提高该算法的寻优效率及求解精度,将该算法用于高光谱影像波段选择中,并与其他常用智能算法进行对比分析,验证本方法的优越性。

## 2 改进斑鬣狗算法

### 2.1 SHO 算法

斑鬣狗是一种复杂、聪明、高度社会化的大型食肉动物,它们生活在非洲和亚洲的热带草原、亚荒漠草原中,通常成群生活和捕猎,当它们发现食物时,会发出声音、动作等信号,并依赖多名同伴构成的交际网络,对猎物发起围捕和进攻,这种方法使得斑鬣狗群体有很高的狩猎成功率。SHO 算法模拟合作捕食机制的四个过程,包括搜索、包围、狩猎和攻击猎物,依据四个过程建立数学模型实现目标优化,基本原理<sup>[12]</sup>如下:

#### 1) 包围猎物

斑鬣狗能够判断目标猎物的位置,不断接近并包围猎物。用数学模型表示这一过程,即认为在当前搜索空间中最佳斑鬣狗确定猎物或目标后,其他

的搜索个体(斑鬣狗)会更新自己的位置,具体描述如下:

$$D_h = |B \cdot P_p(t) - P(t)|, \quad (1)$$

$$P(t+1) = P_p(t) - E \cdot D_h, \quad (2)$$

式中: $D_h$ 表示斑鬣狗个体与猎物之间的距离; $P_p$ 为假定的猎物位置; $P$ 为斑鬣狗个体位置; $t$ 为迭代次数; $B$ 和 $E$ 为系数向量,其计算公式为

$$B = 2 \cdot r_{d1}, \quad (3)$$

$$E = 2h \cdot r_{d2} - h, \quad (4)$$

$$|h| = 5 - \left[ I_{\text{inter}} \left( \frac{5}{M_{\text{inter}}} \right) \right], \quad (5)$$

式中:随机数 $r_{d1}$ 和 $r_{d2}$ 取值范围为 $[0, 1]$ ;  $M_{\text{inter}}$ 为最大迭代次数; $I_{\text{inter}} = 1, 2, 3, \dots, M_{\text{inter}}$ ;控制因子 $h$ 从5到0线性递减。

### 2) 追捕猎物

斑鬣狗通常依靠种群中可信赖的交际网路和本身熟识猎物位置的能力来追捕猎物。假设最佳斑鬣狗个体知道猎物的位置并无限接近猎物,即可认为最佳斑鬣狗所在位置就是猎物位置,其他斑鬣狗个体组成一个集群向着最佳斑鬣狗位置移动,保留当前更新的最好值。在该机制中,由以下方程表示:

$$D_h = |B \cdot P_p - P_k|, \quad (6)$$

$$P_k = P_h - E \cdot D_h, \quad (7)$$

式中: $P_h$ 定义为第一只位置最好的斑鬣狗; $P_k$ 定义为其他的斑鬣狗位置。

$$C_h = P_k + P_{k+1} + \dots + P_{k+N}, \quad (8)$$

$$N = C_{\text{nos}}(P_h, P_{h+1}, P_{h+2}, \dots, P_h + M), \quad (9)$$

式中: $N$ 为斑鬣狗数量; $C_h$ 是 $N$ 个最优解的集合; $M$ 是 $[0.5, 1]$ 之间的随机数; $C_{\text{nos}}$ 是加上 $M$ 之后所有候选解的数量。

### 3) 攻击猎物(局部搜索)

在斑鬣狗对猎物攻击的过程中, $h$ 值不断减少, $E$ 值也随之变化,当 $|E| < 1$ 时,斑鬣狗对猎物发起攻击,反之则是搜索猎物阶段,攻击过程的公式为

$$P(t+1) = \frac{C_h}{N}, \quad (10)$$

式中, $P(t+1)$ 表示更新的最优解的位置,且根据该位置更新其他斑鬣狗的位置。

### 4) 搜索猎物(全局搜索)

斑鬣狗大多分布在最优斑鬣狗集群 $C_h$ 中,它们彼此分散寻找并攻击猎物,利用 $|E|$ 大于1或小于1来强制斑鬣狗远离猎物,扩大搜索阶段以寻找到最优的猎物位置,以保证全局搜索的实现。

## 2.2 黄金正弦算法

黄金正弦算法(golden sine algorithm, Golden-SA)是Tanyildizi<sup>[16]</sup>于2017年提出的一种新的求解优化问题的元启发式算法,它是通过正弦函数来实现的,每个维度的分布都是均匀的。Golden-SA算法的核心是其位置更新方式,首先随机产生 $s$ 个个体的位置,然后假设优化问题的每个解对应搜索空间中相应个体的位置 $X_i^{(t)}$ ,即 $X_i^{(t)} = (X_{i,1}, X_{i,2}, \dots, X_{i,d})$ ,其表示 $d$ 维个体空间中第 $t(t=1, 2, \dots, t_{\text{max}})$ 次迭代后第 $i(i=1, 2, \dots, s)$ 个个体的位置,在 $t+1$ 次迭代中,第 $i$ 个个体的位置更新公式为

$$X_i^{(t+1)} = X_i^{(t)} \sin R_1 + R_2 \sin R_1 \cdot |x_1 P_i^{(t)} - x_2 X_i^{(t)}|, \quad (11)$$

式中: $P_i^{(t)} = (P_{i,1}, P_{i,2}, \dots, P_{i,d})$ 表示第 $t$ 次迭代中 $i$ 个个体的最优位置; $R_1$ 和 $R_2$ 均为随机数, $R_1 \in [0, 2\pi]$ ,  $R_2 \in [0, \pi]$ ,前者决定下一次迭代中个体的移动距离,后者决定下一次迭代中个体的位置更新方向; $x_1$ 和 $x_2$ 是有关于黄金分割数的系数,能够缩小搜索空间使个体更有机会趋近最优值,黄金分割数 $\tau = (\sqrt{5} - 1) / 2$ ,其中, $x_1 = -\pi + 2\pi(1 - \tau)$ ,  $x_2 = -\pi + 2\pi\tau$ 。黄金正弦算法原理图如图1所示。

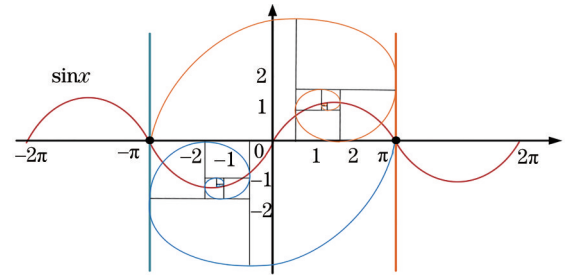


图1 黄金正弦算法原理图

Fig. 1 Principle diagram of Golden-SA

由SHO算法位置更新式(6)、(7)可知,斑鬣狗个体一直以最佳斑鬣狗位置为移动导向,通过不断缩短两者之间的直线距离进行位置更新,相互之间缺乏信息交流,且仅通过线性变化的参数向量 $E$ 控制斑鬣狗扩大搜索范围,并不能更好地模拟出实际的搜捕过程,在处理复杂寻优问题时会陷入局部最优,使得寻优精度较低。本文利用黄金正弦位置更新方法对斑鬣狗算法进行改进,观察式(11)可知,引入黄金正弦策略仅需要提供斑鬣狗个体位置 $P_h$ 及最优个体(猎物)位置 $P_k$ ,而这些信息在SHO算法中可以轻易得到,改进SHO算法的位置更新公式为



$$\mathbf{P}_{h_i}^{(t+1)} = \mathbf{P}_{h_i}^{(t)} \left| \sin R_1 \right| + R_2 \sin R_1 \cdot \left| x_1 \mathbf{P}_{k_i}^{(t)} - x_2 \mathbf{P}_{h_i}^{(t)} \right|, \quad (12)$$

式中： $\mathbf{P}_{h_i}$ 为最佳斑鬣狗位置(猎物位置)； $\mathbf{P}_{k_i}$ 为其他斑鬣狗个体位置。每一次迭代斑鬣狗个体都会与最优个体(猎物)进行信息交流，充分吸收本身与最优个体(猎物)的位置差信息。依据正弦函数与单位圆的关系，算法可以遍历正弦函数上的所有值，即单位圆上的所有点，具有更强的全局搜索能力。利用  $R_1$  和  $R_2$  中的黄金分割系数不仅可以控制位置更新的距离和方向，还可以不断缩小解空间，使得斑鬣狗对能够产生优质解的区域充分探索，更快达到最佳位置，加快了算法的收敛速度，提高了寻优效率，获得更加理想的寻优结果<sup>[17]</sup>。

### 2.3 混沌初始化策略

混沌运动是一种常见的非线性现象，它有着非常强大的规律性、敏感性和遍历性，混沌优化算法被提出后，学者们将其应用到工程、医学等众多领域。随后，学者们又把重点放在了混沌算法的随机性和多样性特点上，将其引入到各类智能优化算法中用于改进算法的寻优性能<sup>[18]</sup>。斑鬣狗算法在寻优过程中，种群序列不断地迭代更新以无限接近最优解，显然，初始种群对后续的寻优效率、精度有很大的影响，而原始算法仅通过随机方式初始化种群，得到的种群不均匀，无法覆盖整个搜索空间。

因此，本文利用混沌映射的方法对 SHO 算法进行初始化，以提高初始斑鬣狗种群的多样性和随机性。最常见的混沌映射函数是 Logistic 函数，但其产生的混沌序列全部分布在  $(0, 0.1)$  和  $(0.9, 1)$  内，存在两端多、中间少的问题，不能均匀分布在整个搜索空间，由此得到的初始种群分布不均匀，影响算法的寻优精度，而 Tent 序列分布均匀且收敛速度更快<sup>[19]</sup>，因此本文采用 Tent 函数初始化种群，其数学表达式为

$$\mathbf{P}_{t+1} = \begin{cases} 2\mathbf{P}_t & 0 \leq \mathbf{P}_t < 0.5 \\ 2(1 - \mathbf{P}_t) & 0.5 \leq \mathbf{P}_t \leq 1 \end{cases}, \quad (13)$$

式中， $\mathbf{P}_t$ 为混沌序列，初始序列  $\mathbf{P}_t = 0.6$ 。

经过伯努利移位(Bernoulli shift)变换后可表示为

$$\mathbf{P}_{t+1} = (2\mathbf{P}_t) \bmod 1. \quad (14)$$

### 2.4 锦标赛选择策略

锦标赛选择策略是一种基于局部竞争选拔机制的方法，该方法通过类似车轮比赛的方式选取适应度值更好的个体，能够让每个个体之间进行充分的交流，使算法更容易跳出局部最优，得到更高的

寻优精度。基于此，将其作用于斑鬣狗算法种群个体适应度值选择上，以改进算法易陷入局部最优问题<sup>[20]</sup>。具体步骤如下：

- 1) 从斑鬣狗种群中随机选择  $k$  个个体，通常  $k = 2$ 。
- 2) 比较每个个体的适应度值，选择适应度最好的个体将其赋予 1 分。
- 3) 重复上述步骤，选择分数前  $n$  (最优子集个数) 个个体组成新一代最优种群。

该方法保证了每个个体均按照此策略进行比较，使个体能够充分发挥自身的寻优能力。特别是，该方式利用个体之间的适应度值关系进行比较，而不是适应度值本身，避免了超级个体的存在，从而影响算法过早收敛。而个体得分越高被选择的概率就会越大，意味着该个体更加优秀，相比较斑鬣狗算法通过贪婪选择方法得到的最优种群，更加接近最优目标，并获得更高的寻优精度。

## 3 结合黄金正弦和混沌斑鬣狗算法的高光谱波段选择

高光谱遥感图像分类是根据光谱信息和地物信息将图像中的每一个像素点进行分类，对于含有几十甚至几百个波段的高光谱遥感数据而言，相邻波段间相关性较大，含有大量冗余信息，给图像分类带来很大困扰，因此，利用波段选择方法进行数据降维十分重要。每个波段都可看成是一个特征，那么波段选择又可称为特征选择，本文提出结合黄金正弦和混沌斑鬣狗算法(GSSHO)的高光谱波段选择方法，以达到数据降维、提高分类精度的目的，总体方法框图如图 2 所示。

### 3.1 二进制编码

波段选择是一个难以用十进制编码直接求解的离散组合优化问题，需要对其进行二进制编码，把个体特征的真实值或转换值切换为逻辑值“0”或“1”。GSSHO 算法的编码长度与原始高光谱数据维数相同，每个波段有两个候选条件，即被选择和不被选择，通过二进制编码很容易映射，“1”表示当前波段被选择用于分类，“0”表示当前波段被丢弃。本文二进制编码转换由 Sigmoid 函数计算得到，公式为

$$\text{Sigmoid}(\mathbf{P}) = 1 / [1 + \exp(-\mathbf{P})], \quad (15)$$

$$\mathbf{P}_b = \begin{cases} 1, \text{rand}(\ ) < \text{Sigmoid}(\mathbf{P}) \\ 0, \text{rand}(\ ) \geq \text{Sigmoid}(\mathbf{P}) \end{cases}, \quad (16)$$

式中： $\mathbf{P}$ 为个体位置； $\mathbf{P}_b$ 为最终二进制编码结果；

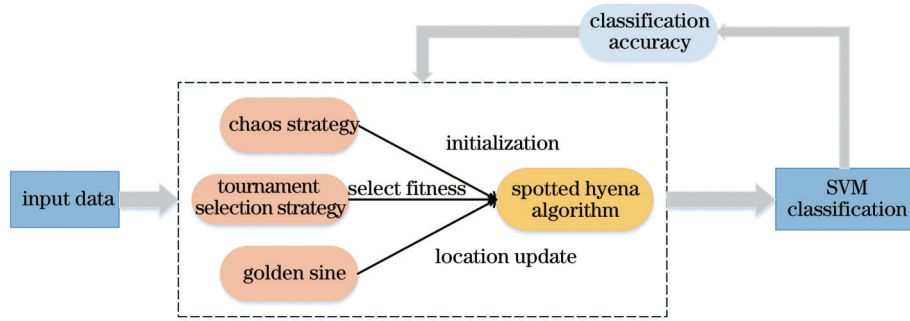


图 2 高光谱波段选择方法框图

Fig. 2 Block diagram of hyperspectral band selection method

rand( )是区间[0, 1]的随机数。

### 3.2 适应度函数

在优化算法中,适应度函数用于评价目标解的质量,构造合适的函数是搜索最优目标的关键步骤。高光谱波段选择的目的是选用尽可能少的波段获得更高的分类精度,本文以最大化分类精度和最小化所选波段数目来构造适应度函数。综合考虑上述两种因素,将最小的适应度函数值设置为最佳值,适应度函数定义为

$$F(i) = \alpha [1 - OA(i)] + \beta \frac{R}{N}, \quad (17)$$

式中: $F(i)$ 为第*i*个个体的适应度值; $R$ 和 $N$ 分别为选择的波段数和波段总数; $\alpha$ 和 $\beta$ 为权重参数,这里取 $\alpha = 0.8, \beta = 1 - \alpha$ ;  $OA(i)$ 为根据所选波段子集进行分类所得的分类精度,可表示为

$$OA(i) = F_p / (F_p + F_N), \quad (18)$$

式中: $F_p$ 为正确分类像元数; $F_N$ 为错误分类像元数。

### 3.3 基于 GSSHO 算法的波段选择

将波段选择优化问题抽象成斑鬃狗群体捕猎过程,算法基本思想为:如果高光谱遥感图像有  $n$  条波段,则构造一个  $n$  维的向量空间,然后设置  $m$  个斑鬃狗个体构成种群,在  $n$  维空间中寻找猎物,以适应度函数为准则不断更新斑鬃狗位置直到得到满意的寻优结果或达到预设迭代次数。算法基本步骤如下:

Step1 Tent 混沌策略初始化斑鬃狗种群  $P_i (i = 1, 2, \dots, n)$ 。

Step2 设置初始参数  $N, R_1, R_2, x_1, x_2$  及最大迭代次数  $M_{inter}$ 。

Step3 计算每个搜索个体的适应度值  $F(i)$ 。

Step4 根据黄金正弦位置更新式(12),在搜索空间中寻找最佳搜索个体。

Step5 定义一组最优解,使用式(8)和式(9)直到找到满意的结果。

Step6 使用式(10)更新搜索个体的位置。

Step7 检查在给定搜索空间中是否有搜索个体越界,并对其进行调整。

Step8 计算更新搜索个体的适应度值,利用锦标赛策略选出更好的解,更新  $P_h$ 。

Step9 更新斑鬃狗种群  $C_h$ ,并更新搜索个体的适应度值。

Step10 如果满足停止条件,则算法停止。否则,返回步骤 5。

Step11 满足停止条件后,返回得到的最优解。

结合 GSSHO 算法的具体流程图如图 3 所示。

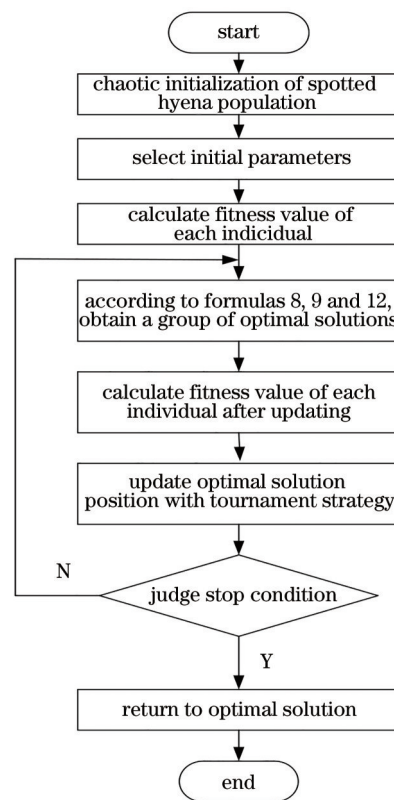


图 3 GSSHO 算法流程图

Fig. 3 Flow chart of GSSHO algorithm

## 4 实验与结果分析

### 4.1 实验数据集

采用两个高光谱遥感影像数据集进行实验,一个是 Salinas 高光谱遥感数据,是由 AVIRIS 成像光谱仪所拍摄的美国加利福尼亚州的 Salinas 山谷遥感影像,去除水汽吸收严重的波段后,共有 204 个波段,总共分为 16 类。另一个是 Pavia Centre 高光谱遥感数据,是由 ROSIS 成像仪对意大利的帕维亚城所成的影像数据,剔除噪声影响的波段,一般使用剩下的 102 个光谱波段所成的图像,尺寸为 1096 pixel × 715 pixel,共有 9 类。分别以 2:8、5:5、8:2 的比例划分训练集和测试集,多次实验均值结果表明,三种划分方式得到的分类精度几乎相同,而训练集样本数越多,所用时间越长,因此考虑时间成本,以 2:8 的比例划分数据集,即将数据的 20% 作为训练集,80% 作为测试集。

### 4.2 实验结果与分析

实验在 3 GHz CPU 和 4 GB RAM 的 Windows 7 系统, Matlab 2018b 开发环境中实现。为了验证 GSSHO 算法用于波段选择的有效性,采用支持向量机(SVM)对选择后的特征进行分类<sup>[21]</sup>,并将该算法与粒子群算法(PSO)、模拟退火算法(SA)、遗传算法(GA)和灰狼算法(GWO)进行对比实验。每个智能优化算法设置的种群大小均为 10,终止迭代次数  $M_{\text{iter}}$  为 50,且各算法均独立重复实验 30 次,取均值为最终结果。算法具体参数设置如表 1 所示。

表 2 各算法对 Salinas 数据集分类结果

Table 2 Classification results of Salinas data set by each algorithm

Band selection algorithm	Overall classification accuracy / %	Number of bands	Fitness function value	Kappa coefficient	Running time / s
GSSHO	93.32	32	0.0836	0.9283	10852
PSO	92.73	40	0.0974	0.9201	12360
SA	93.17	92	0.1372	0.9259	19863
GA	93.64	121	0.1619	0.9285	8979
GWO	93.08	94	0.1366	0.9241	10593

由表 2 和表 3 显示的测试结果表明, GSSHO 算法的总体性能明显优于其他优化算法。相比较 PSO、SA 和 GWO 算法, GSSHO 算法能够选出更少的波段数,并且得到更高的分类精度和 Kappa 系数,虽然 GA 算法的分类精度和 Kappa 系数略高于 GSSHO 算法,但其所选波段个数是 GSSHO 算法波段个数的 4~5 倍, GSSHO 算法显示出了更强大的

表 1 实验所涉及的算法参数

Table 1 Algorithm parameters involved in experiments

Algorithm	Parameter	Set value
GSSHO	Control factor $h$	[0, 5]
	Number of candidate solutions $M$	[0.5, 1]
	Individual $k$	2
PSO	Weight $\omega$	[0.2, 0.9]
	Learning factors $c_1, c_2$	1.49
	Speed $v$	[0, 6]
SA	Annealing coefficient $\alpha_T$	0.99
	Initial temperature $T_0$	0.1
GA	Cross probability $C$	0.4
	Variation probability $M$	0.4
GWO	Synergy coefficient $a$	[0, 2]

以适应度函数值(Fitness)为主要评价指标,并通过总体分类精度(OA)、所选波段个数(N)、Kappa 系数和运行时间综合评价 5 种算法的性能。其中,总体分类精度是被正确分类的像素数与总数之比,精度越高分类效果越好;Kappa 系数用于判断分类模型结果和实际分类效果的一致性,其值越大越好;各优化算法均以适应度函数来评价解的质量,它直接影响了算法的收敛能力和能否找到最优解,同时,也是综合评价分类精度和降维效果的重要指标,适应度值越小表示该算法性能越好。将 GSSHO 算法与其他算法进行对比实验, Salinas 数据集和 Pavia Centre 数据集的各测试结果如表 2 和表 3 所示。

数据降维能力。特别的是, GSSHO 算法所选出的波段数分别为 32 条和 11 条,对于原高光谱数据集 204 条和 102 条波段来说,显著降低了数据维度,有效剔除了冗余信息,使后续高光谱图像分类更加准确高效,最大程度地减轻了“维数灾难”的影响。对于 Pavia Centre 数据集,其分类精度达到了 99.08%,得到了非常好的分类效果。



表 3 各算法对 Pavia Centre 数据集分类结果  
Table 3 Classification results of Pavia Centre data set by each algorithm

Band selection algorithm	Overall classification accuracy / %	Number of bands	Fitness function value	Kappa coefficient	Running time / s
GSSHO	99.08	11	0.0315	0.9856	7038
PSO	98.57	15	0.0396	0.9787	7485
SA	99.02	37	0.0797	0.9849	8096
GA	99.14	58	0.1259	0.9868	4980
GWO	98.96	45	0.0932	0.9833	6975

同时, GSSHO 算法与其他算法相比, 迭代结束后能够获得更小的适应度值, 远小于 GA 算法, 且两个数据集的适应度值均趋近于 0, 表明该算法具有更好的收敛性, 不仅体现了算法能够得到更好的分类效果, 还能展现其强大的数据降维能力。由于黄金正弦位置更新方式的引入加强了 SHO 算法个体的全局遍历能力, 所提出的适应度函数公式对所需解决问题具有针对性, 在算法迭代过程中能够更加

准确地判断出优秀个体, 充分体现了算法优越的探索和寻优能力。但是, GSSHO 算法引入了混沌初始化和锦标赛策略, 算法较为复杂, 虽然得到了更好的降维效果和分类精度, 但是算法运行时间较长、效率较低, 在时间效率方面还有待提升。为了更加直观地表现算法的分类效果, 两个数据集由 5 种算法所得的波段子集的图像分类结果分别如图 4 和图 5 所示。

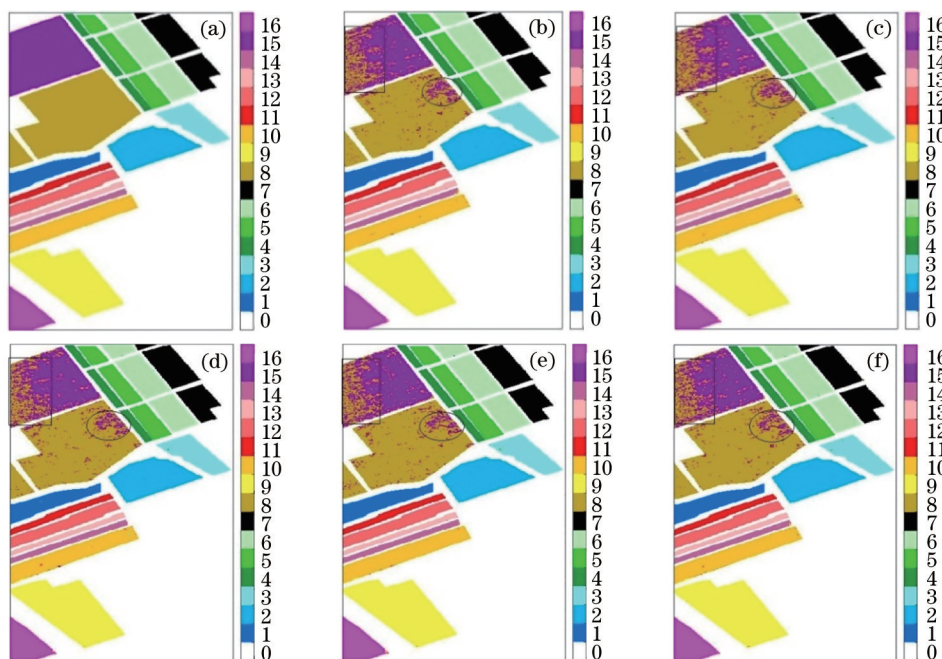


图 4 Salinas 数据集分类图。(a) 原始图像; (b) GSSHO 算法; (c) PSO 算法; (d) SA 算法; (e) GA 算法; (f) GWO 算法  
Fig. 4 Salinas data set classification chart. (a) Original image; (b) GSSHO algorithm; (c) PSO algorithm; (d) SA algorithm; (e) GA algorithm; (f) GWO algorithm

从图 4 可以清晰地看出, 类别 8 (Grapes\_untrained) 和类别 15 (Vinyard\_untrained) 的分类误差较大, 5 种算法对于这两类的分类精度如表 4 所示, 除 GA 算法外, GSSHO 算法对两类别的分类精度均高于其他算法, 充分验证了其优越性。很明显, 两类别易被 SVM 分类器混淆, 尤其是类别 15 分类精度较低, 说明两类别的光谱特性较为接近, 波

段选择算法选出的波段子集所含信息不够理想, 不足以清晰地区分两类, 还需要后续不断地深入研究。从图 5 可知, Pavia Centre 数据集各算法分类图像与原始图像非常接近, 各算法分类精度接近 1, 分类效果十分理想。

为了进一步验证 GSSHO 算法在不同分类算法上的表现, 获得更加可信的分类结果, 使用 4 种分类

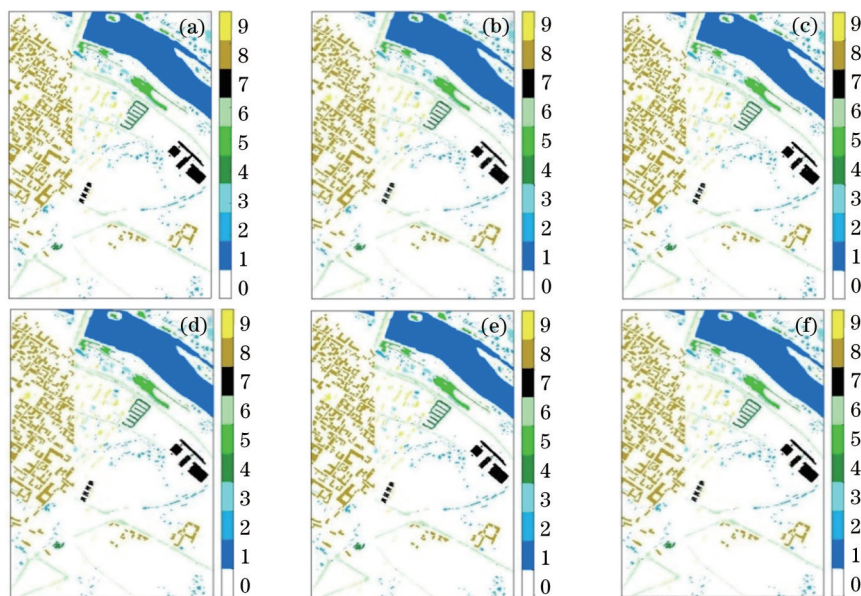


图 5 Pavia Centre 数据集分类图。(a)原始图像;(b) GSSHO 算法;(c) PSO 算法;(d) SA 算法;(e) GA 算法;(f) GWO 算法  
 Fig. 5 Pavia Centre data set classification chart. (a) Original image; (b) GSSHO algorithm; (c) PSO algorithm; (d) SA algorithm; (e) GA algorithm; (f) GWO algorithm

表 4 各算法对两类别分类精度

Table 4 Accuracy of each algorithm for two kinds of classification

Algorithm	GSSHO	PSO	SA	GA	GWO
Grapes_untrained	90.38%	89.06%	90.03%	90.55%	89.85%
Vinyard_untrained	68.72%	64.89%	68.28%	68.56%	67.56%

算法: SVM、K 近邻(KNN)<sup>[22]</sup>、决策树(DT)<sup>[23]</sup>、朴素贝叶斯(NB)<sup>[24]</sup>, 分别来衡量波段选择后的子集优劣, 并采用上述 OA、N、Fitness 三个指标来评估算法分类器性能, 两个数据集的实验结果如表 5 和表 6 所示。结果表明, GSSHO 算法在各分类算法上都能得到良好的分类精度和降维效果, 特别是, SVM 分类器始终保持最高的总体分类精度、最小的

波段个数和适用度值, 而 NB 算法表现最不理想, 适应度值高达 0.2183, GSSHO 算法表现的收敛性能和寻优效果不佳。这是因为 SVM 非常适用于小样本、非线性数据, 尤其是对于高维数据分类有着优异的表现, 而 NB 算法在处理特征较多、相关性较大的数据时, 分类效果不佳。

表 5 Salinas 数据集各分类算法结果

Table 5 Results of different classification algorithms in Salinas data set

Classification algorithm	Overall classification accuracy /%	Number of bands	Fitness function value
SVM	93.32	32	0.0836
KNN	90.45	43	0.1197
DT	87.91	41	0.1336
NB	77.07	35	0.2183

表 6 Pavia Centre 数据集各分类算法结果

Table 6 Results of different classification algorithms in Pavia Centre data set

Classification algorithm	Overall classification accuracy /%	Number of bands	Fitness function value
SVM	99.08	11	0.0315
KNN	98.18	39	0.0942
DT	97.19	21	0.0627
NB	88.71	15	0.1111



综上所述, GSSHO 算法在波段选择问题上表现出了良好的收敛性和寻优能力, 能够更好地达到提高分类精度和减少数据冗余的目的, 且具有普遍适用性, 可以用于多个分类算法。但是, 在今后的研究中, 还需要不断克服运行时间较长的缺点, 提高 GSSHO 的优化能力, 进一步减少波段的选择数量, 尤其是重点关注较易混淆类别的分类问题, 不断探索能够代表各类别的特征波段或波段子集, 进一步提升高光谱图像分类精度。

## 5 结 论

提出的 GSSHO 算法不仅保留了原始 SHO 算法较强的开发和搜索能力, 还充分利用了黄金正弦算法的全局遍历能力和混沌策略的多样性特点, 同时, 综合考虑了分类精度和波段个数, 基于二者构造出了新型的适应度函数。在 Salinas 数据集和 Pavia Centre 数据集上进行实验, 依据所提 GSSHO 算法实现波段子集的生成和优化, 并对高光谱影像像素进行分类, 所选波段个数分别为 32 和 11, 分类精度高达 93.32% 和 99.08%。将各项分类结果与 PSO、SA、GA、GWO 4 种算法进行比较, 虽然 GSSHO 算法分类精度略低于 GA 算法, 但相比 GA 算法所选出的 121 条和 58 条波段, 其降维效果显著, 且其他各项指标均明显好于 GA 算法。同时, 为证明 GSSHO 算法具有普遍适用性, 将其用于 SVM、KNN、DT、NB 4 种分类算法中, 结果表明该算法在 4 种分类器中都能得到较好的分类结果, 其中, 在 SVM 分类器中表现最佳。总之, GSSHO 算法能够更加有效地识别出信息量更大的波段子集, 得到更高的分类精度和更少的波段个数, 具有令人满意和稳定的性能, 非常适用于高光谱波段选择。

## 参 考 文 献

- [1] Chutia D, Bhattacharyya D K, Sarma K K, et al. Hyperspectral remote sensing classifications: a perspective survey[J]. Transactions in GIS, 2016, 20(4): 463-490.
- [2] 黄冬梅, 张晓桐, 张明华, 等. 考虑空间相关性的半监督局部保持投影的高光谱图像特征提取[J]. 激光与光电子学进展, 2019, 56(2): 021003.  
Huang D M, Zhang X T, Zhang M H, et al. Feature extraction of hyperspectral images based on semi-supervised locality preserving projection with spatial-correlation[J]. Laser & Optoelectronics Progress, 2019, 56(2): 021003.
- [3] 马世欣, 刘春桐, 李洪才, 等. 基于线性嵌入和张量流形的高光谱特征提取[J]. 光学学报, 2019, 39(4): 0412001.  
Ma S X, Liu C T, Li H C, et al. Feature extraction based on linear embedding and tensor manifold for hyperspectral image[J]. Acta Optica Sinica, 2019, 39(4): 0412001.
- [4] 关世豪, 杨枕, 卢珊, 等. 基于注意力机制的多目标优化高光谱波段选择[J]. 光学学报, 2020, 40(21): 2128002.  
Guan S H, Yang G, Lu S, et al. Multi-objective optimization of hyperspectral band selection based on attention mechanism[J]. Acta Optica Sinica, 2020, 40(21): 2128002.
- [5] Zhang A Z, Ma P, Liu S H, et al. Hyperspectral band selection using crossover-based gravitational search algorithm[J]. IET Image Processing, 2019, 13(2): 280-286.
- [6] Wang M W, Wu C M, Chen M L, et al. A band selection approach based on Lévy sine cosine algorithm and alternative distribution for hyperspectral image[J]. International Journal of Remote Sensing, 2020, 41(9): 3429-3445.
- [7] 王立国, 魏芳洁. 结合遗传算法和蚁群算法的高光谱图像波段选择[J]. 中国图象图形学报, 2013, 18(2): 235-242.  
Wang L G, Wei F J. Band selection for hyperspectral imagery based on combination of genetic algorithm and ant colony algorithm[J]. Journal of Image and Graphics, 2013, 18(2): 235-242.
- [8] Li Y J, Liang K, Tang X J, et al. Waveband selection based feature extraction using genetic algorithm[C]//2017 IEEE 4th International Conference on Cyber Security and Cloud Computing, June 26-28, 2017, New York, NY, USA. New York: IEEE Press, 2017: 223-227.
- [9] Chang Y L, Liu J N, Chen Y L, et al. Hyperspectral band selection based on parallel particle swarm optimization and impurity function band prioritization schemes[J]. Journal of Applied Remote Sensing, 2014, 8: 084798.
- [10] Xu M X, Shi J Q, Chen W, et al. A band selection method for hyperspectral image based on particle swarm optimization algorithm with dynamic subswarms[J]. Journal of Signal Processing Systems, 2018, 90(8/9): 1269-1279.
- [11] 胡梦林, 万幼川, 王明威, 等. 基于混沌杜鹃搜索算法的高光谱影像波段选择[J]. 微电子学与计算机,

- 2018, 35(4): 124-129.
- Hu M L, Wan Y C, Wang M W, et al. Band selection based on chaotic cuckoo search algorithm for hyperspectral image[J]. *Microelectronics & Computer*, 2018, 35(4): 124-129.
- [12] Dhiman G, Kumar V. Spotted hyena optimizer: a novel bio-inspired based metaheuristic technique for engineering applications[J]. *Advances in Engineering Software*, 2017, 114: 48-70.
- [13] Dhiman G, Kumar V. Spotted hyena optimizer for solving complex and non-linear constrained engineering problems[M]//Yadav N, Yadav A, Bansal J C, et al. *Harmony search and nature inspired optimization algorithms. Advances in intelligent systems and computing*. Singapore: Springer, 2019, 741: 857-867.
- [14] 李杰. 斑点鬣狗优化算法及其应用研究[D]. 南宁: 广西民族大学, 2019: 34-41.
- Li J. Spotted hyena optimizer and its applications[D]. Nanning: Guangxi University for Nationalities, 2019: 34-41.
- [15] 贾鹤鸣, 姜子超, 李瑶, 等. 基于模拟退火斑点鬣狗优化算法的特征选择[J]. *应用科技*, 2020, 47(1): 74-79.
- Jia H M, Jiang Z C, Li Y, et al. Feature selection based on simulated annealing spotted hyena optimization algorithm[J]. *Applied Science and Technology*, 2020, 47(1): 74-79.
- [16] Tanyildizi E, Demir G. Golden sine algorithm: a novel math-inspired algorithm[J]. *Advances in Electrical and Computer Engineering*, 2017, 17(2): 71-78.
- [17] 周新, 邹海. 融合黄金正弦混合变异的自适应樽海鞘群算法[J]. *计算机工程与应用*, 2021, 57(12): 75-85.
- Zhou X, Zou H. Adaptive salp swarm algorithm with golden sine algorithm and hybrid mutation[J]. *Computer Engineering and Applications*, 2021, 57(12): 75-85.
- [18] Chen W J, Xiao Y C. An improved ABC algorithm and its application in bearing fault diagnosis with EEMD[J]. *Algorithms*, 2019, 12(4): 72.
- [19] Rather S A, Bala P S. Swarm-based chaotic gravitational search algorithm for solving mechanical engineering design problems[J]. *World Journal of Engineering*, 2020, 17(1): 97-114.
- [20] 张振兴, 杨任农, 房育寰, 等. 自适应 Tent 混沌搜索的蚁狮优化算法[J]. *哈尔滨工业大学学报*, 2018, 50(5): 152-159.
- Zhang Z X, Yang R N, Fang Y H, et al. Ant lion optimization algorithm based on self-adaptive Tent chaos search[J]. *Journal of Harbin Institute of Technology*, 2018, 50(5): 152-159.
- [21] 方晓玉, 李晓斌, 郭震. 一种改进的混合灰狼优化支持向量机预测算法及应用[J]. *激光与光电子学进展*, 2020, 57(12): 122801.
- Fang X Y, Li X B, Guo Z. Improved hybrid grey wolf optimization support vector machine prediction algorithm and its application[J]. *Laser & Optoelectronics Progress*, 2020, 57(12): 122801.
- [22] 于挺, 杨军. 基于 K 近邻卷积神经网络的点云模型识别与分类[J]. *激光与光电子学进展*, 2020, 57(10): 101510.
- Yu T, Yang J. Point cloud model recognition and classification based on K-nearest neighbor convolutional neural network[J]. *Laser & Optoelectronics Progress*, 2020, 57(10): 101510.
- [23] 尚朝轩, 王品, 韩壮志, 等. 基于类决策树分类的特征层融合识别算法[J]. *控制与决策*, 2016, 31(6): 1009-1014.
- Shang C X, Wang P, Han Z Z, et al. Feature-level fusion recognition algorithm based on analogy decision tree classification[J]. *Control and Decision*, 2016, 31(6): 1009-1014.
- [24] Su H S. Multi-source fuzzy information fusion method based on Bayesian optimal classifier[J]. *Acta Automatica Sinica*, 2008, 34(3): 282-287.