Blind Separation of Mixed Speech Based on an Improved Firefly Optimization Algorithm (Postprint)

Authors: Li Zhucheng, Huang Xianglin

Date: 2018-06-19T00:00:00+00:00

Abstract

The Firefly Optimization Algorithm (FOA) is developed by simulating the characteristic of fireflies in nature communicating with each other through luciferin, and demonstrates high effectiveness in searching for global extremum points of complex functions. To address the limitations of traditional blind source separation optimization algorithms that significantly impact separation performance, this paper proposes a mixed speech blind separation algorithm based on improved firefly optimization. The new algorithm transforms the flight span of fireflies from a fixed value to adaptive adjustment via a newly constructed function, thereby accelerating convergence speed while simultaneously avoiding premature convergence. Experimental results demonstrate that, compared with blind separation algorithms based on gradient, standard firefly, and particle swarm optimization, the proposed algorithm achieves superior separation performance for mixed speech signals, with improvements in both convergence speed and separation accuracy.

Full Text

Preamble

Blind Separation of Speech Mixtures Based on Improved Glowworm Swarm Optimization

Li Zhucheng[1,2], Huang Xianglin[1]† 1. Faculty of Science & Technology, Communication University of China, Beijing 100024, China 2. Business College, Beijing Union University, Beijing 100025, China

Abstract: Glowworm Swarm Optimization (GSO) is a population-based metaheuristic inspired by the luminescent communication behavior of natural fireflies. GSO demonstrates high effectiveness in searching for global extrema of complex functions. Addressing the limitations of traditional optimization algorithms for Blind Source Separation (BSS), this paper proposes a blind speech separation algorithm based on an Improved GSO (IGSO). The new algorithm adaptively adjusts the flight span through a newly constructed function, which accelerates convergence while preventing premature convergence. Experimental results demonstrate that compared with BSS algorithms based on gradient descent, standard GSO, and Particle Swarm Optimization (PSO), the proposed algorithm achieves superior separation performance for mixed speech signals, with improvements in both convergence speed and separation accuracy.

Key words: glowworm swarm optimization; blind source separation; blind speech separation; flight span

0 Introduction

Swarm intelligence optimization algorithms represent a new research branch in artificial intelligence. These algorithms do not require continuity or convexity of objective functions and constraints, offering unique advantages in solving complex problems with characteristics of low cost, high speed, and robustness. They have been applied across numerous fields including computational mathematics, aerospace, transportation, computer science, electronic communications, power systems, material mechanics, and ecosystem modeling. Blind Source Separation (BSS) refers to the technique of recovering independent source signals from observed mixtures without prior knowledge of the source signals or transmission channel characteristics. BSS represents a current research hotspot in signal processing, with applications spanning speech recognition, image processing, biomedical engineering, and mechanical fault detection.

Generally, BSS algorithms consist of two essential components: an objective function and an optimization algorithm. The objective function specifies or describes the statistical independence of recovered signals, while the optimization algorithm seeks the optimal solution (separation matrix) in the objective function space. The performance of BSS algorithms—particularly convergence speed and separation accuracy—depends primarily on the latter, making the selection of an appropriate optimization algorithm a key challenge in BSS technology. Traditional BSS optimization algorithms are gradient-based; unless initial parameter values are properly chosen, these algorithms yield suboptimal solutions. However, due to the blind nature of the problem, obtaining reasonable parameter values is extremely difficult. Moreover, when the objective function space is discontinuous or non-differentiable, such algorithms become inapplicable.

To overcome the limitations of traditional optimization algorithms and leverage the advantages of swarm intelligence methods, this paper investigates the use of an improved glowworm swarm optimization algorithm to solve the mixed speech blind separation problem. Glowworm Swarm Optimization is a novel swarm intelligence algorithm inspired by the collective behavior of fireflies communicating through luminescence, demonstrating good performance in multimodal function optimization. Two primary versions of firefly algorithms exist: Glowworm Swarm Optimization (GSO), proposed by Indian scholars Krishnanand et al., and Firefly Algorithm (FA), proposed by Cambridge scholar Yang. Both share similar bionic principles but differ in implementation details, with GSO being more commonly adopted. The following research is based on GSO.

1.1 Mathematical Model

Consider a set of N-dimensional source signals $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T \in \mathbb{R}^N$, an $M \times N$ mixing matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{M\times N}$, and a set of M-dimensional observed signals $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T \in \mathbb{R}^M$. The linear instantaneous mixing model can be expressed as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)$$

where $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_M(t)]^T \in \mathbb{R}^M$ represents noise, which is generally neglected. The BSS task involves determining a separation matrix $\mathbf{W}$ to recover source signals from observations:

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t)$$

where $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_N(t)]^T \in \mathbb{R}^N$ represents the separated signals—estimates of the source signal vector. The case $M = N$ is termed determined BSS, $M < N$ is underdetermined BSS, and $M > N$ is overdetermined BSS.

1.2 Signal Preprocessing

To simplify algorithm complexity and improve computational efficiency, observed signals $\mathbf{x}(t)$ typically require preprocessing before separation.

a) Centering: This involves centralizing the observed signals to ensure zero-mean vectors. The most direct method subtracts the mean: $\mathbf{x} \leftarrow \mathbf{x} - E[\mathbf{x}]$.

b) Whitening: Also known as sphering, this transforms the observed signals using a whitening matrix $\mathbf{V}$ such that $E[\mathbf{v}\mathbf{v}^T] = \mathbf{I}$, making the components uncorrelated. Whitening methods include Principal Component Analysis (PCA), Singular Value Decomposition (SVD), and autocorrelation function-based approaches. Whitening reduces computational load by nearly half, significantly improving efficiency.

c) Objective Function Selection: Negentropy and kurtosis are common measures of statistical independence in signal processing. However, kurtosis is extremely sensitive to outliers, often leading to unstable results, making negentropy more widely applied. Yet negentropy is difficult to compute accurately, prompting literature [14] to propose an approximation formula. Analysis reveals that under certain conditions, a mapping relationship exists between negentropy and kurtosis.

2.1 Algorithm Principle

The algorithm initially distributes multiple glowworm individuals randomly in the solution space. Each glowworm represents a solution in the objective function space and possesses a corresponding luciferin value at any given moment, where the luciferin magnitude reflects the quality of its position (i.e., its fitness). Each glowworm maintains a sensing range determined by its decision radius. Within this range, it identifies neighbors with higher luciferin values, then moves toward selected neighbors based on movement probabilities. The decision radius is subsequently updated. Through iterative repetition, glowworms congregate around brighter individuals, effectively locating multiple local or global extrema of the objective function. The GSO algorithm comprises five main steps: luciferin update, neighbor selection, movement probability calculation, position update, and decision radius update.

2.2 Mathematical Description

2.2.1 Luciferin Update

Higher luciferin values confer stronger attractiveness, increasing the probability of other individuals moving toward the glowworm. The luciferin value relates to the fitness function $Fitness(\mathbf{x}_i(t))$ as:

$$l_i(t+1) = (1-\rho)l_i(t) + \gamma \cdot Fitness(\mathbf{x}_i(t))$$

where $l_i(t)$ represents the luciferin value of glowworm i, ρ denotes the luciferin decay factor, γ denotes the luciferin enhancement coefficient, and $\mathbf{x}_i(t) \in \mathbb{R}^M$ represents the position of individual i.

2.2.2 Neighbor Selection

Glowworm i identifies neighbors with higher luciferin values within its sensing range to form neighbor set $N_i(t)$:

$$N_i(t) = \{j : d_{ij}(t) < r_d^i(t); l_i(t) < l_j(t)\}$$

where $d_{ij}(t)$ is the distance between glowworms i and j, and $r_d^i(t)$ represents the decision radius of individual i.

2.2.3 Movement Probability Calculation

Based on the determined neighbor set $N_i(t)$, the roulette probability $p_{ij}(t)$ is calculated:

$$p_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} (l_k(t) - l_i(t))}$$

2.2.4 Position Update

According to the movement probability $p_{ij}(t)$, glowworm i updates its position:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + s \cdot \frac{\mathbf{x}_j(t) - \mathbf{x}_i(t)}{\|\mathbf{x}_j(t) - \mathbf{x}_i(t)\|}$$

where s represents the flight span of glowworms.

2.2.5 Decision Radius Update

Glowworm i updates its decision radius based on neighbor set $N_i(t)$:

$$r_d^i(t+1) = \min\{r_s, \max\{0, r_d^i(t) + \beta(n_t - |N_i(t)|)\}\}$$

where β is a regulation factor, r_s denotes the normal sensing radius, and n_t represents the neighborhood threshold.

2.3 Flight Span Function Strategy

Although standard GSO incorporates adaptive decision radii—where the radius shrinks when many high-luciferin individuals are present and expands otherwise, effectively avoiding local optima—it cannot resolve the inherent conflict between fixed flight span values and the trade-off between convergence speed and separation accuracy. This paper proposes a functional transformation of the flight span parameter to enable real-time adaptive adjustment: larger values enhance global search capability and accelerate convergence during early iterations, while smaller values ensure local search precision in later stages. The flight span function is expressed as:

$$s(p, n) = \psi \cdot e^{-\mu p/n} + \xi$$

where ψ, μ, and ξ are positive constants, and ξ represents the flight span threshold.

3 Blind Separation Based on Improved GSO

3.1 Algorithm Flowchart

Figure 1 [Figure 1: see original paper] illustrates the blind separation process using the improved glowworm swarm optimization algorithm.

3.2 Algorithm Implementation Steps

Based on Figure 1, the algorithm implementation proceeds as follows:

a) Read observed signals $\mathbf{x}(t)$ and perform centering and whitening.

b) Initialize parameters ρ, γ, β, s, n_t, and l_0. Randomly generate n separation matrices as the glowworm swarm, initializing each glowworm's position $\mathbf{x}_i(0)$ and decision radius $r_d^i(0)$ in the search space. Compute initial fitness $Fitness(\mathbf{x}_i(0))$ using negentropy (scaled proportionally) as the fitness function.

c) Determine the swarm' s optimal position $\mathbf{x}_{opt}$ and best fitness $Fitness_{opt}$ from the current results.

d) Update luciferin values $l_i(t)$ using equation (6), compute neighbor sets $N_i(t)$ using equation (7), and calculate movement probabilities $p_{ij}(t)$ toward neighbors using equation (8).

e) Update position $\mathbf{x}_i(t+1)$, decision radius $r_d^i(t+1)$, and flight span $s(p,n)$ for the next iteration using equations (9)-(11).

f) Compare fitness values to obtain the maximum. If this value surpasses $Fitness_{opt}$, update the optimal position $\mathbf{x}_{opt}$ and best fitness $Fitness_{opt}$.

g) If the separation accuracy meets requirements or termination conditions are satisfied, output $\mathbf{x}_{opt}$ and $Fitness_{opt}$ and stop; otherwise, return to step **d)**.

3.3 Computer Simulation Experiments

MATLAB experiments compare the proposed algorithm with BSS methods based on gradient descent, standard GSO, and PSO.

3.3.1 Experimental Environment and Parameter Settings Three speech analog signals were collected via microphone, digitized, and sampled at 20,000 points as source signals (Figure 2 [Figure 2: see original paper]), with histograms shown in Figure 3 [Figure 3: see original paper]. Mixed signals generated through random matrix mixing and their histograms are displayed in Figures 4 [Figure 4: see original paper] and 5 [Figure 5: see original paper], respectively. Algorithm parameters are listed in Table 1 .

Table 1 Algorithm Parameter Values

Parameter	Value
ρ	0.4
γ	0.6
β	0.08
n_t	5
l_0	5
r_s	3
n	50

Parameter	Value
ψ	0.5
μ	0.2
ξ	0.2

3.3.2 Experimental Results and Analysis **a)** Figure 6 [Figure 6: see original paper] shows separated signals after 30 iterations of the proposed algorithm (histograms in Figure 7 [Figure 7: see original paper]). Comparison with Figure 2 reveals that source signals are well-recovered, while other algorithms have not yet converged at this stage.

b) Figure 8 [Figure 8: see original paper] depicts the two-dimensional trajectory of two separation matrix variables $W_{1,1}$ and $W_{2,1}$, visually illustrating the glowworm swarm' s search trajectory toward the optimal solution.

c) Figure 9 [Figure 9: see original paper] compares fitness curves across algorithms, demonstrating that the proposed algorithm outperforms gradient-based, standard GSO, and PSO-based BSS methods in both convergence speed ($speed_{Gradient} < speed_{PSO} < speed_{GSO} < speed_{IGSO}$) and precision ($prec_{Gradient} < prec_{PSO} < prec_{GSO} < prec_{IGSO}$).

d) Figure 10 [Figure 10: see original paper] presents scatter plots of source and recovered signals, showing that after convergence, some separated signals exhibit changed order and phase relative to source signals—a characteristic uncertainty inherent to BSS [15].

e) Parameter Analysis: The reference values in Table 1 are suitable for most applications. Parameter ρ represents the luciferin decay factor; a value of 1 indicates no memory. γ reflects the fitness extraction ratio, determining the incremental luciferin update per iteration. β, the neighborhood change rate, should not be too large to avoid limiting neighbor range to bounds. The flight span s should not be fixed but functionally transformed. The neighbor threshold n_t requires careful selection: too small values reduce sample diversity and may erroneously discard excellent individuals, while too large values slow convergence. Initial luciferin l_0 can be arbitrary. The decision radius r_s must balance global awareness and local search capability. Population size n, ψ, and μ lack stable reference values and should be tuned according to the objective function landscape.

4 Conclusion

To enhance the global search performance of glowworm swarm optimization, this paper first functionalizes the flight span parameter and then proposes a mixed speech blind separation algorithm based on improved GSO to address limitations of traditional BSS optimization methods. Comparative experiments with gradient-based, standard GSO, and PSO-based BSS algorithms demonstrate that the proposed algorithm achieves superior separation performance with improvements in both convergence speed and accuracy.

References

[1] Ou Xudong, Zhang Tianqi, Yan Zhenhua, et al. Blind convolution speech separation based on multi-band ordering [J//OL]. Application Research of Computers, 2016, 33 (5): 1481-1485.

[2] Abbas N A. Image encryption based on independent component analysis and Arnold' s cat map [J]. Egyptian Informatics Journal, 2016, 17 (1): 139-146.

[3] Jadhav M S H, Dhang M D N. Extraction of fetal ECG from abdominal recordings combining BSS-ICA & WT Techniques [J]. International Journal of Engineering, 2017, 10 (1): 869-875.

[4] Farhat M, Gritli Y, Benrejeb M. Fast-ICA for mechanical fault detection identification in electromechanical systems for wind turbine applications [J]. International Journal of Advanced Computer Science and Applications, 2017, 8 (7): Article ID 080759.

[5] Shan Rong, Zhao Zengshun, Chen Panfei, et al. Network modeling and assessment of ecosystem health by a multi-population swarm optimized neural network ensemble [J]. Applied Sciences, 2016, 6 (6): Article ID 175.

[6] Mavaddaty S, Ebrahimzadeh A. Blind signals separation with genetic algorithm and particle swarm optimization based on mutual information [J]. Radioelectronics and Communications Systems, 2011, 54 (6): 315-324.

[7] Li Qiaoyan, Quan Haiyan. Analyze gravity tide signal based on ICA with PSO [C]// Proc of IEEE ICSPCC. Piscataway, NJ: IEEE Press, 2015: 1-4.

[8] Zhang Nian, Liu Tianyou. The application of ant colony optimization algorithm in linear-combination blind source separation problem [C]// Proc of IEEE CISP. Piscataway, NJ: IEEE Press, 2009: 1-4.

[9] Krishnanand K N, Ghose D. Detection of multiple source locations using a glowworm metaphor with applications to collective robotics [C]// Proc of IEEE SIS. Piscataway, NJ: IEEE Press, 2005: 84-91.

[10] Krishnanand K N, Ghose D. Glowworm swarm optimization for multimodal search spaces [M]// Panigrahi B K, Shi Y, Lim M H. Handbook of Swarm Intelligence. Concepts, Principles and Applications. Berlin: Springer-Verlag, 2011: 451-467.

[11] Kaipa K N, Ghose D. Glowworm swarm optimization: theory, algorithms, and applications [M]. Cham, Switzerland: Springer Nature, 2017.

[12] Yang Xinshe. Firefly algorithms for multimodal optimization [C]// Proc of the 5th International Conference on Stochastic Algorithms: Foundations and

Applications. Berlin: Springer-Verlag, 2009: 169-178.

[13] Yang Xinshe. Multiobjective firefly algorithm for continuous optimization [J]. Engineering with Computers, 2013, 29 (2): Article s00366-012-0254-1.

[14] Comon P. Independent component analysis, a new concept? [J]. Signal Processing, 1994, 36 (3): 287-314.

[15] Liu Ju. The theory and application of blind signal processing [M]. Beijing: Science Press, 2013.

Note: Figure translations are in progress. See original paper for figures.