

SRP-PHAT Source Location Algorithm Based on Chaos Artificial Bee Colony Algorithm

Yecai GUO^{1,2,a}, Jiping WU¹, Sainan ZHU¹

¹School of Electronic & Information Engineering, Nanjing University of Information Science and Technology, Nanjing, 210044, China

²Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment, Nanjing, 210044, China

^aemail: guo-yecai@163.com

Keywords: Chaotic Mapping; Artificial Colony; SRP-PHAT; Sound Source Localization Algorithm

Abstract. Aiming at defects of large computational complexity and poor real-time performance of the traditional steering response power-phase transform (SRP-PHAT) sound source localization algorithm, a SRP-PHAT sound source localization algorithm based on chaos artificial bee colony algorithm was proposed. In this proposed algorithm, the beam power function of microphone array is used as the objective cost function to establish the mathematical model of source localization, the global coverage of chaos and the fast convergence speed and strong robustness of the artificial colony algorithm are combined to realize three dimensional sound source localization. The experimental results show that the proposed algorithm has lower computational complex and high positioning accuracy, and it can satisfy the demand of the actual positioning system.

Introduction

In recent years, the sound source location technology based on microphone array has become one of the research hot spots in the field of speech signal processing[1][2][3][4]. Microphone array positioning algorithm can obtain the location of the sound source in video conferencing systems, intelligent robots, and listening devices, etc. In the actual environment, the background noise and the reverberation of the room have a serious impact on the performance of the sound source location algorithm. However, under the condition of strong noise and reverberation, the sound source location algorithm based on steering response power and phase transform(SRP-PHAT) is of good robustness and accuracy, which has achieved widespread attention and research[5].

The sound source location algorithm based on SRP-PHAT adopts the weighted phase transformation method to compute the steering response power(SRP) of signals and to find the maximum in the target area, which is the estimated position of the target area. However, the SRP space has many local maximums[6][7][8]. To find a global maximum, the computationally-intensive grid-search methods are used, whereas the grid search is too expensive for a real-time system[9][10]. To reduce the amount of calculation, some improved SRP-PHAT algorithms were proposed, however, their robustness is bad or the array shape is limited. Since the energy function of the SRP-PHAT algorithm has the characteristics of many peaks in the indoor environment, a SRP-PHAT source location algorithm based on chaos artificial bee colony algorithm, which combines with the advantages of the artificial colony algorithm and the overall characteristic of the chaotic thoughts, is proposed.

Array Signal Model

In the actual environment, the speech received by the microphone array is always influenced by the background noise and reverberation. And assume that the sound source signal is $s(t)$, then the signal received by the i th microphone can be expressed as

$$x_i(t) = s(t) * h_i(t) + v_i(t), i = 1, \dots, M \quad (1)$$

where $h_i(t)$ is the channel impulse response, which describes the channel characteristics between

the signal and the i th microphone, and it is highly dependent on the sound source position and the position of the corresponding microphone. $v_i(t)$ is the background noise, which is independent from the sound source signals.

SRP-PHAT Sound Source Location Algorithm

The SRP-PHAT sound source location algorithm is a combination of the short analysis feature of the SRP sound source localization algorithm and the insensitive property of phase transform (PHAT) to the environment. So in the actual complex environment, it can effectively restrain the effects of the reverberation and noise, and obtain sound source position estimation.

Then, its output power of the beam can be expressed as

$$E(\mathbf{q}) = \sum_{k=1}^M \sum_{l=k+1}^M \int_{-\infty}^{+\infty} \Psi_{kl}(\omega) X_k(\omega) X_l^*(\omega) e^{j\omega(\tau_l - \tau_k)} d\omega \quad (2)$$

where

$$\tau_l = \frac{\|\mathbf{q}_l - \mathbf{b}\|}{c} \quad (3)$$

where \mathbf{b} is the position vector of the imaginary source coordinates, M is the number of the microphones in microphone array, $X_k(\omega)$ is the Fourier transform of the k th microphone signal $x_k(t)$ after adding window, τ_l is the direct time between sound source and the l th microphone, the symbol $\|\cdot\|$ is 2-norm of the vector, \mathbf{q}_l is the position of the l th microphone, c is the sound speed, $\Psi_{kl}(\omega)$ is the weighted function, which can be expressed as

$$\Psi_{kl}(\omega) = \frac{1}{|X_k(\omega) X_l(\omega)|} \quad (4)$$

Obviously, the beam power output function is function of the target space point coordinates. Therefore, the evaluation of the beam power output function is defined as fe (function evaluation), then the beam power output function is computed via using a point coordinate calculation within the target space, i.e., fe is computed once.

The SRP-PHAT sound source location algorithm can find the position, namely, the estimated location of the object sound source, corresponding to the maximum of the SRP via searching the predetermined space point by point. And it can be expressed as

$$\hat{\mathbf{q}}_s = \arg \max_{\mathbf{q} \in \mathcal{Q}} \{E(\mathbf{q})\} \quad (5)$$

where \mathcal{Q} is the preset space.

In a real environment, due to the presence of the room multipath reflection wave, the strong background noise and the limited microphone array shape, the beam power output function $E(\mathbf{q})$ will appear more peak phenomenon in some positions of the predetermined space, and it makes the global maximum search extremely complex. The traditional SRP-PHAT sound source location algorithm adopts the grid method to search the peak, which divides the space into the grid and calculates the beam output power function of each grid point within the target area. In order to ensure a high positioning accuracy, it needs to mesh very close, and there are a lot of grid points, so the total computational load is very large. Thus, it is difficult for the SRP-PHAT sound source localization algorithm to be applied to real-time processing and the practical application.

SRP-PHAT Sound Source Location Algorithm Based on Chaos ABC Algorithm

Artificial bee colony algorithm. Artificial bee colony (Artificial bee colony, ABC) algorithm is based on the evolution in a random and objective way of group consisting of candidate solutions to obtain the optimal solution of the multidimensional functions and multi-peak functions. And it has the advantages of strong flexibility and robustness, which isn't restricted by domain knowledge.

In the artificial colony algorithm, each bee's position represents a feasible solution to the problem, the quality of the bee represents the quality of corresponding feasible solution, and the swarm is composed of employed bees, onlooker bees, and scout bees. The process that bees find the best source is a process to solving parameter optimization iteration, and in each iteration, the employed bees search locally higher-quality honey nectar in the corresponding area of the position depending on the honey nectar. In all honey nectars, high-quality honey nectar will attract more the onlooker bees to honey, whereas the onlooker bees search the honey in the neighborhoods. When they find high quality honey nectar, they will notify corresponding employed bees to update the honey nectar, otherwise, they will be replaced with the scout bees at the same time. If the honey nectars aren't updated for several consecutive iterations, they will be given up and a new one will be randomly generated to search again. In the iterative optimization process of the ABC algorithm, the employed bees always can obtain the good solution, the onlooker bees can improve the convergence speed, and the scout bees can effectively avoid local optimal solution.

Chaotic thoughts. Chaos is a kind of nonlinear phenomenon existing in the nature, has strong randomness, ergodicity, and can product unrepeatabe random points within certain limits of time and range. If given enough time and space, the optimal solution of the problem must be able to be found based on the idea of chaos. Chaos Logistic map is a kind of typical chaotic mapping and can produce a series of the unpredictable sequence, and its mathematical formula can be expressed as

$$x_{i+1} = \mu x_i (1 - x_i) \quad (6)$$

where μ is a Logistic parameters in $[0,4]$, x_0 is a random point in $[0,1]$. This map is in a state of chaos when $x \in [0,1], x \neq 0.25, 0.5, 0.75$ and $\mu \in [3.57, 4]$ and can produce the random point with aperiodic.

SRP-PHAT source location algorithm based on chaos ABC algorithm. ABC algorithm with fast convergence rate, less adjusting parameters, and strong robustness, and so on, can't cover the entire solution space because the initial nectar and new nectar in stage of scout bee are randomly produced. While the chaos with very strong ergodicity and randomness can randomly and repeatedly search each solution in a certain space range and is of strong overall. Therefore, on the basis of the SRP-PHAT sound source localization algorithm, chaotic thoughts and artificial colony algorithm are introduced to propose a SRP-PHAT sound source location algorithm based on chaos artificial bee colony algorithm (known to it as CABC-SRP).

Assume that the size of the colony is NS , the size of the employed bees is NE , and the size of the onlooker bees is NU . The minimum and maximum of the boundary of the three-dimensional space are q_{\min} and q_{\max} , respectively, and they are all 3D space vector. $q^{(0)}$ is used as the initial position of the nectar, $q^{(k)}$ is used as the k th generation nectar. $search_i$ denotes the searching times that the employed bees search at the position of the i th nectar, $maxCycle$ represents the maximum number of the iterations. The implementation steps of the proposed algorithm are given as follows:

Initialization of the chaos ABC(CABC). Within the scope of the predetermined space, chaos factor $Z_0(z_1, z_2, z_3)$ is randomly produced and nectars $(q_1, q_2, \dots, q_{NE})$ with NE individual are also produced, i.e.,

$$q_i = q_{\min} + Z_i(q_{\max} - q_{\min}) \quad (7)$$

$$Z_i = [z_{i1}, z_{i2}, z_{i3}], \quad i = 1, 2, \dots, NE \quad (8)$$

where z_{it} is the chaotic factor for the i th iterative in the t th position, and it can be expressed as

$$z_{it} = \mu z_{(i-1)t} (1 - z_{(i-1)t}), \quad t = 1, 2, 3 \quad (9)$$

It can produce NE new nectar $q(0)$, which is on behalf of the estimates of the sound source location.

(2) To calculate the quality of the nectar. According to Eq.(2), the quality of the initial nectar can be calculated to obtain the fitness value and it can be expressed as

$$f_i = E(\mathbf{q}_i) \quad (10)$$

where f_i is the fitness of the i th nectar.

(3) The onlooker bees search new nectar in the areas. For the k th generation nectar $\mathbf{q}(k)$, new location within the neighborhood of every nectar is searched, i.e.,

$$\mathbf{n}_{it} = \mathbf{q}_{it} + \varphi_{it} \times (\mathbf{q}_{it} - \mathbf{q}_{mt}) \quad (11)$$

where $i, m = 1, 2, \dots, NE$ and $i \neq m$. n_{ij} is the t th element of \mathbf{n}_j , and $t \in \{1, 2, 3\}$. φ_i is a random number between $[-1, 1]$.

(4) The onlooker bees update nectar. The quality $E(\mathbf{n}_i)$ of the nectar within the neighborhood is calculated and compared with f_i to keep the higher quality nectar, then the searching times is updated according to $search_i = search_i + 1$.

(5) The employed bees choose nectar. The employed bees choose good nectar according to the information provided by the onlooker bees, and the choice probability is expressed as

$$p(\mathbf{q}_i) = \frac{f_i}{\sum_{j=1}^{NE} f_j} \quad (12)$$

(6) The employed bees update nectar. The employed bees select the current nectar and search nectar within the neighborhood according to the step (3) and calculate the quality of the nectar. If the nectar quality is higher, the nectar is updated and $search_i = 0$. Otherwise, $search_i = search_i + 1$.

(7) The k th generation of the best nectar and the corresponding fitness value are recorded. the k th generation nectar quality of NE individual is compared with the best nectar of the $(k-1)$ th generation to choose the highest quality nectar as the best nectar of the k th generation, namely, the k th iteration optimization of the sound source location estimate.

(8) Scout bees investigate. If $search_i > limit$, the best nectar will be given up, the scout bees randomly select new nectar according to Step (1), and it is replaced with the onlooker bee.

(9) Judge whether maximum number of iterations is reached or not. if any, search process is over and the position of the best nectar is obtained and is used as estimates of the target sound source location. If no, go back to Step (2).

Computer Simulation

In order to verify the effectiveness of the localization algorithms, computer simulation experiments with simulating indoor environment were carried out. Room size is $3 \text{ m} * 4 \text{ m} * 4 \text{ m}$, Environmental parameters $SNR = 15 \text{ dB}$, $RT60 = 400 \text{ ms}$, microphone array is the square array and shown in Fig.1.

In the Fig.1, microphone array is $3 * 3$ square planar array, the array element number M was set to 9, the distance d between the adjacent microphone was set to 0.4 m , the microphone array was close to YOZ metope and at the center of the metope. Sound source signal is a girl reading in English recorded in advance, and its sampling frequency was 25 kHz , the frame length was 512 points, the moving frame was about 50%.

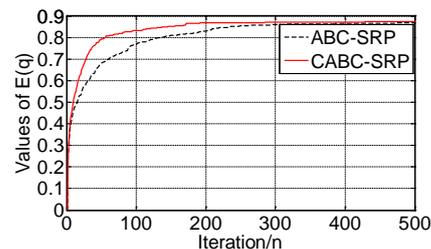
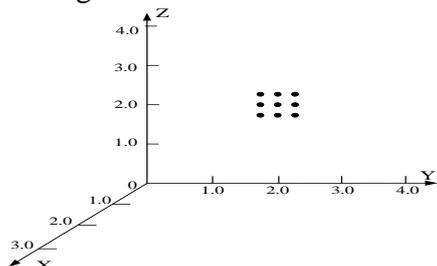


Fig.1. Microphone array indoor model Fig.2. Convergence curves of different localization algorithms

Convergence performance. In order to verify the convergence of the positioning algorithm, the function value $E(q)$ and RMSE(Root Mean Square Error) were used as evaluation indexes. The higher the Function value $E(q)$ and the smaller the RMSE, the better the performance of the positioning algorithm.

The sound source was fixed at $s(1.0, 2.0, 1.0)$, the SRP - PAHT sound source localization algorithm based on artificial bee colony algorithm (ABC-SRP) and CABC-SRP were simulated and compared. The simulation parameters were given as follows: NP = 40, and limit = 50. The simulation results were shown in Fig.2 and Table 1.

Table 1 The positioning results of different localization algorithms

Localization algorithms	RMSE/m	Iteration/n	positioning results		
			x/m	y/m	z/m
ABC-SRP	0.0307	300	0.9896	2.0084	2.9724
CABC-SRP	0.0192	180	1.0159	2.0082	2.9927

The Table 1 and Fig.2 shows that the ABC-SRP algorithm and the CABC-SRP algorithm can realize 3D indoor sound source localization. At the same time, under the condition of the same parameter, the CABC-SRP has a faster convergence speed and a higher positioning accuracy.

Positioning performance. In order to further verify the performance of the positioning algorithm, the RMSE, TNC(The Number of Calls),and RAC(Relative Amount of Calculation) were used as evaluation standard. Among them, the TNC denotes the number of calling fe function, the greater the TNC, the greater the amount of calculation. RAC is the relative ratio of TNC and the called number used by grid method and can be expressed as $RAC = TNC / N_{grid}$. Assume that the sound source location accuracy based on the grid method is 1cm, and the number of calling fe function by grid method is $N_{grid} = 300 \times 400 \times 400 = 4.8 \times 10^7$.

When sound source was fixed in $(1.0, 2.0, 3.0)$, $(1.0, 1.0, 1.0)$, and $(1.0, 2.0, 1.0)$, respectively, the ABC-SRP algorithm, the CABC-SRP algorithm, and the SRP-PHAT source Localization algorithm using Stochastic Region Contraction(SRC-SRP) were simulated and compared. Among them, for the ABC-SRP algorithm, the parameters NP = 40, limit = 50, maxCycle = 300. For the CABC-SRP algorithm, the parameters NP = 40, limit = 50, maxCycle = 220. For the SRC-SRP algorithm, the parameters J0 = 8000 and npoints = 100. The positioning results were shown in Table 2.

Table 2 The positioning results of different localization algorithm

Algorithms	RMSE/m	TNC/n	RAC/%
SRC-SRP	0.0390	13609	0.028
ABC-SRP	0.0299	12146	0.025
CABC-SRP	0.0297	8910	0.019

The Table 2 shows that, in terms of the SRC-SRP algorithm, the ABC-SRP algorithm and the CABC-SRP algorithm proposed by this paper have a faster positioning speed, and their positioning accuracy have an improvement of about 1cm. Especially, the positioning speed of the CABC-SRP algorithm has an improvement of about 35%, so it has better real-time performance.

Conclusions

The SRP-PHAT sound source localization algorithm has a good anti-jamming performance, but it is necessary to search the whole target space and has a large amount of calculation and poor real-time performance. To overcome the defects of the SRP-PHAT sound source localization algorithm, the global coverage of chaos and the search mechanism of the artificial bee colony algorithm are introduced into the SRP-PHAT localization algorithm to propose SRP-PHAT source location algorithm based on chaos artificial bee colony algorithm. Compared with the localization algorithm in the literature[9], the proposed algorithm has a less amount of calculation, a faster

positioning speed, and a higher positioning accuracy. Accordingly, it is more suitable for positioning in a actual system.

Acknowledgements

Project supported the Major Project of Nature Science Foundation of Higher Education Institution of Jiangsu Province, China(Grant No.13KJA510001), Jiangsu Scientific Research Achievements in Industrialization Project, China(JHB 2012-9), and Jiangsu Provincial College Students' innovative training program(201310300051Y)

References

- [1] Salih Y, Sebastian P, Yap V V. Video localization using array of microphones[C]. Intelligent and Advanced Systems (ICIAS), 2010: 1-4.
- [2] Tourbabin, V.; Rafaely, B., Theoretical framework for the optimization of microphone array configuration for humanoid robot audition[J]. Audio, Speech, and Language Processing, IEEE/ACM Transactions on, 2014, 22(12):1803-1814.
- [3] Scharrer R., Vorlander M., Sound field classification in small microphone arrays using spatial coherences[J], Audio, Speech, and Language Processing, IEEE Transactions on, 2013, 21(9):1891-1899.
- [4] Dmochowski J P, Benesty J, Affes S. A generalized steered response power method for computationally viable source localization[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2007, 15(8) : 2510-2526.
- [5] Cobos M. Marti, Amparo. A modified srp-phat functional for robust real-time sound source localization with scalable spatial sampling[J], Signal Processing Letters, IEEE, 2011, 18(1):71-74.
- [6] Hoang Do, Silverman, H.F., Ying Yu. A real-time srp-phat source location implementation using stochastic region contraction(src) on a large-aperture microphone array[C], IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2007:15-20.
- [7] Qin Quande, Cheng Shi, Li li, et al. Artificial Bee colony algorithm: a Survey[J]. CAAI Transactions on Intelligent Systems, 2014, 9(2):127-135.
- [8] Sharma, T.K., Pant, M., Bansal, J.C. Some modifications to enhance the performance of Artificial Bee Colony[C], IEEE Congress on Evolutionary Computation (CEC), 2012:10-15.
- [9] Fan Jiu-lun,ZHANG Xue-feng. Piecewise Logistic Chaotic Map and Its Performance Analysis [J]. Acta Electronica Sinica, 2009,37(4):720-725.
- [10] Wang Kai, Zong Zhiya,Sun Xiaowei, et al. Study on improved adaptive eigenvalue decomposition algorithm for acoustic source localization[J]. Chinese Journal of Scientific Instrument, 2013, 34(6):42-47.