Improved Ant Colony Algorithm of Image Retrieval Methods

Hualin Sun

1Changzhou Institute of mechatronic technology, Changzhou 213164, China.
sunhualin021@126.com

Abstract

Ant colony algorithm since it has a better convergence and parallelism, widely used in the data retrieval, however, is not high, due to the characteristics of retrieval object use seriously affect the accuracy of retrieval, and according to this problem, this paper proposed an improved ant colony algorithm, this algorithm will retrieve objects comprehensive characteristics into the ant colony algorithm, and solve the convergence speed and computational complexity of the algorithm, obtained good results in image retrieval.

Keywords: Ant colony algorithm; Image retrieval; Convergence speed; Computational complexity

1. Introduction

As a result of image retrieval in pattern recognition, image analysis, image understanding, and other technologies play a very important role, it is considered one of the key technologies of digital image processing, the attention of researchers and engineers. However, the nature of the image is often complicated, in order to accurately extraction conforms to human visual effect of target and area, it is still a particularly challenging task [1-2].

Many scholars focus on the research in the field of image retrieval, makes the retrieval method based on different ideas inspired emerged, which in addition to some of the classic algorithms such as retrieval method based on threshold method, edge detection and region tracking retrieval method, the researcher puts forward a new method combined with the new mathematical algorithm. But any kind of image retrieval method using must to specific research and application of the premise, not the only, standards, retrieval method, all did not adapt to the optimal search criteria, after reading the literature in the field of image retrieval in recent years, found that many scholars devoted to today will introduce a new mathematical method of image retrieval, such as wavelet analysis, fuzzy mathematics, fractal theory have the specific application in image retrieval, and researchers have come to realize that to get better search results, the effective combination of become a trend, which combines the method of fuzzy set theory and artificial intelligence retrieval technology play a new role, and gradually been applied in the field of image retrieval, made the image retrieval method has been further expanded.

Ant colony algorithm is a kind of heuristic random search algorithm of robust characteristics, it puts forward the basic model, further improve the minimization problem, knapsack problem and applied to the data, and in job shop scheduling problems in practical application of engineering projects have achieved good effect. At present, as the scholars at home and abroad to further study and application of ant colony algorithm has extended to more widely in the field of technology, such as image processing, pattern recognition, etc. By the development of ant colony algorithm and its application process it is easy to see that human creative thinking will bring a bigger to ant colony algorithm application prospect. Ant colony algorithm in image processing applications are still relatively new, the theoretical research and books are very few, just see in the literature of
recent years some scholars have research in this field. From theoretical analysis, image processing the problem is that the data processing, and has the characteristics of various complex, discretization of pixel data, and the ant colony algorithm suitable for solution of this problem, it is used for image processing, both from the time efficiency and processing effect will bring certain progress, so the ant colony algorithm for image processing algorithm is one of the research is very meaningful.

2. Related Works

2.1. Ant Colony Algorithm

Bionic optimization algorithm is an important branch in the field of artificial intelligence, including biological, natural selection and genetic mechanism of genetic algorithm and ant colony foraging behavior of ant colony algorithm and simulated birds group feeding behavior of the particle swarm algorithm, etc. Ant colony algorithm is the genetic algorithm and neural network algorithm, after the heuristic search algorithm and a heuristic search algorithm applied to combinatorial optimization problems, relative to other kinds of heuristic search algorithm, ant colony algorithm has obvious advantages, its successful application aroused people's great interest to ant colony algorithm, and attracted a number of researchers engaged in the research of ant colony algorithm. Ant colony algorithm not only has the characteristics of intelligent search, global optimization, but also has the stability, the positive feedback, distributed computing and easily combined with other algorithms, etc. Distributed computing makes the algorithm easy to parallel implementation, and ongoing communication between individuals and pass, is helpful to find a better solution, avoid falling into local optimum. Due to the strong robustness, on the basis of the basic ant colony algorithm model modification can be applied to other problems. Therefore, ant colony algorithm was developed to solve complex optimization problems in many areas provides a powerful tool, has the very broad development prospects [4-6].

However, ant colony algorithm is still at the primary research stage, from the ant colony algorithm of related academic papers published, most applications also only stay in the simulation phase, and the research is mostly practical problems in the experimental conditions or to simplify the constraint conditions under the premise of, especially in the dynamic optimization, stochastic optimization problem, multi-objective optimization problems still need further deepening; People from the algorithm model, according to different problems many different types of improved model is put forward, but the ant colony algorithm of the model is not used widely, and the basic ant colony algorithm model and cannot be directly applied to solve the problem of the optimization of concrete, for different problems in the reality, for each of the different problems to design a corresponding model is not realistic, so should further explore the high-level unified mechanism of ant colony algorithm model, combined with the study of complex systems, proposed a more general model of ant colony algorithm; From the point of view of theoretical analysis, The basic characteristics of ant colony algorithm is as follows:

(1) the basic ant colony algorithm is a system. Nature of the ant colony has three basic characteristics of the system, namely, diversity, relevance and integrity, so as to constitute a system. As for abstract ant colony foraging behavior of ant colony algorithm, if the algorithm itself as a whole, will find that it has the characteristics of the system, and this is the bionic optimization algorithm is one of the most important characteristics. In the basic ant colony algorithm, the use of multiple ant result is much better than that of solving the single ant solving result, so it don't have to add and sex, the whole is greater than the sum of the parts, so the ant colony algorithm is a system.

(2) distributed computing. Through the analysis can be found, the nature of real ant colony behavior that is distributed, when the ant colony need to complete a job, many of
these ants for the same job for a common purpose, and the final task completion will not be affected because of the defects of some individuals. By the same token, the ant colony algorithm also reflected the distributed group behavior and each artificial ants start at the same time in the problem space point of multiple tectonic problems independently of each other, and the whole solution of the problem won't because of some artificial ants cannot successful solutions and affected. The basic ant colony algorithm is regarded as a distributed multi-agent system, it independently with the multipoint problem space and solution search, not only makes the algorithm has strong global search ability and also increase the reliability of the algorithm.

(3) the self organization. This is another important feature of basic ant colony algorithm, and it is also the genetic algorithm, particle swarm optimization, artificial neural network and artificial immune algorithm, there are characteristics of the bionic optimization algorithm. In the abstract sense, since the organization is under no outside makes the process of system entropy increase, namely the evolution of the system from disorderly to orderly process. When algorithm to start early, single artificial ant disorderly looking for solutions, after a period of evolution algorithm, artificial ants tend to look for the more and more close to some of the solution of the optimal solution, it is reflected from disorderly to orderly self-organization evolution. Self-organizing greatly enhance the robustness of the algorithm, which makes the self-organization of ant colony algorithm does not need to treat all aspects of the question are aware, and therefore easy to be applied into the class of problems.

(4) positive feedback. Feedback is an important concept in cybernetics, it represents the adverse effect of information input to output. Positive feedback and negative feedback feedback points two kinds, the former refers to is his behavior to strengthen future behavior, while the latter is now behavior to weaken the future behavior. By the foraging behavior of real ants in the nature can be seen, the ants will eventually find the shortest path, directly depends on the shortest path pheromone accumulation, and the accumulation of pheromone is a positive feedback process. Ants when a path, the more the stronger the legacy in the amount of information on the path, then the greater the ant chooses the path of the probability that all the ants will eventually converge to a shortest path, this constitutes a positive feedback phenomenon. Single positive feedback or negative feedback exists in linear system, is unable to realize the system of self organization, self-organizing system with positive feedback and negative feedback, in order to realize the system to create and update [7-8].

2.2. Image Retrieval

In the study of image and application, people often interested in some parts of the image. These sections are often referred to as the target or prospects, they generally correspond to the specific, unique in the images with area. In order to target recognition and analysis, the need to separate out the relevant area, on the basis of this is likely to increase the use of the target. Image retrieval, is to divide the image into several mutually overlapping area process, each area represents the target and background such as meaningful physical objects. Which regions with a consistent attributes, and attribute has the obvious difference between adjacent area.

For the following classification of image retrieval methods. First of all, for gray image retrieval is often based on pixel gray value of two properties: the discontinuity and similarity. Pixels in the region is generally has a gray similarity, and in the area on the border between general with gray discontinuity. So the retrieval algorithm can be divided into use on the basis of regional gray discontinuity algorithm based on boundary and the use area of gray similarity algorithm based on region. Secondly, according to different treatment strategy in the process of retrieval, the retrieval algorithm and can be divided into parallel and serial algorithm. In parallel algorithm, all judgments and decisions were independently and simultaneously, and in serial algorithm, the result of the early treatment
General serial algorithm needed to calculate for a long time, but with a strong ability to resist noise also. Currently widely used in image retrieval research methods are: 1, the threshold of retrieval method. The method to determine an image gray value range of gray threshold, then using the threshold value compare to each pixel in the image gray level, according to the comparison results to the corresponding pixel classification, so as to achieve the purpose of retrieval. Threshold method based on the characteristics of the image itself can be divided into single threshold retrieval and retrieval method, the threshold value according to the retrieval algorithm all the characteristics or principles, can be divided into the histogram and histogram transformation method, the most kinds of spatial variance method and least error method and homogenization, co-occurrence matrix method, the maximum entropy method, fuzzy set method, the characteristics of spatial clustering method, etc. 2, region growing method. Similiar to the basic idea is to combine qualitative pixel area, which makes use of the local space of the image, which can effectively overcome the disadvantage of other methods of image retrieval space discontinuous, but it will usually cause excessive image retrieval. 3, edge test. The basic characteristics of the image is the edge, it is the result of local characteristics of discrete image. Image edge detection method is to use the first derivative of extreme value or the second derivative of zero information to provide the basic foundation of edge point, and the classical edge detection method is sensitive structure of image grayscale step change difference operator for image retrieval. 4, the retrieval method based on artificial neural network. The basic idea is through training multilayer perceptron to get linear decision function, and then use decision function to classify the pixels to achieve the purpose of retrieval. 5, the method based on fuzzy set theory set theory can better describe the fuzziness and randomness in human vision, because the image edge, corner or in the area is difficult to accurately described, and the relationship between uncertainty and fuzziness, the fuzzy set theory is introduced into image retrieval, can achieve good retrieval effect. And some image retrieval method based on mathematical morphology such as image retrieval technology, retrieval technology based on wavelet analysis and transform, image retrieval method combined with genetic algorithm and so on.

Ant colony algorithm, as a new type of intelligent bionic model also has been widely used in image processing field, because its discreteness and parallelism characteristics with digital image processing is applied, in recent years, the application of ant colony algorithm in image retrieval, image feature extraction, image matching, image texture classification, and other fields have achieved quite rich research achievements, but because of the complexity of the algorithm and image processing model itself, making it the low efficiency in solving the problem of image processing, the calculation time is longer, and have yet to find a kind of widely used in image retrieval of the retrieval model, so there is plenty of research space.

3. Improved Ant Colony Algorithm of Image Retrieval Methods

3.1. Model of the Improved Ant Colony Algorithm

In general background, goal, constitute an image edges and noise, good image features must have typicality and representative, and can distinguish between the content of the image, in order to prepare the way for further image analysis, so the need for feature extraction. Usually occupy an image of target and background for the most part of the image, so the grey value can be used to differentiate between the two kinds of content, and the boundary and noise point of just a handful of images, it is not enough to use grey value to distinguish clearly, through the analysis can be found around the edge points and noise pixel gray value are often larger change, in order to show the change, the gray gradient are introduced to describe the two image content as the second feature of the
image and the boundary pixels and noise is obviously high gradient pixel. Finally, in order to better distinguish boundary point and noise point to introduce a description sim pixel neighborhood characteristics of elements, called the neighborhood characteristic value, the extraction process is as follows:

Take $3 \times 3$ neighborhood of each pixel, established variable $sim=0$, for each pixel in the neighborhood, calculate the grey value $(i, j)$, and then by a given threshold $T$ to compare the difference, if the difference is less than $T$, the $sim=sim+1$ repeat the process until each pixel is get a sim values. Sim is representative of neighborhood pixels of similar degree, for the target and background pixel, pixels usually $sim = 8$, and for edge pixels and noise points, generally edge pixels sim value is bigger than noise. Threshold $T$ here for different images to different values for image smoothing, $T$ value is small, and if an image detail is more, the $T$ value would have been a bigger, from experience, the $T$ value generally take between 50-90.

All the ants in the basic ant colony algorithm when the initial random search, practiced according to certain probability of selection, but the direction is still has certain blindness, thus will cause a lot of invalid search, reduce the efficiency of the algorithm, so the introduction of fuzzy clustering method to specify the clustering center, ants search for guidance, so as to improve the efficiency of algorithm. Clustering center is refers to the limited number of representative and the image in pixels. The clustering center initialization process is as follows:

1. Determine the number of clustering centers and their grey value  
Belonging to the same goal of gray levels of pixels with similar usually, gray histogram can show these pixels gathered around different grey value degree, clustering results often depend on these grey value. So based on the gray-level histogram of the original image, choose $n$ of histogram peak as the clustering center, the $n$ is the number of initial clustering center. So that the original algorithm to calculate the circulation between pixels is transformed to pixel compared with a few peaks, guide to the ant clustering center gathered at the same time, reduced the searching process, reduces the computational complexity.

2. To determine the gradient value of cluster centers  
General goals of inner pixels in the image and the gray gradient value is small, and the majority in the image, the boundary pixels and noise point of gray gradient value is large, usually at the same time, the boundary pixels number is greater than the number of noise pixels, so can be set to those contains more wide pixels and the clustering center of the gray-level gradient value is 0, for the rest of the clustering center, calculate the gradient value:

$$gf = \frac{1}{m} \max\left(\sum_{i=1}^{m} gd(i, j)\right)$$

The $gf$ for value and gradient, $gd(i, j)$ is gradient values of pixels $(i, j)$, $m \times n$ for image size.

3. Determine the clustering center neighborhood characteristic value  
Combining the above analysis, according to different kinds of pixel in the image of neighborhood characteristics, often gradient of 0 clustering centers are the target pixels or background pixels, the neighborhood similarity is bigger, so can set its neighborhood characteristic value of 8, and for those who gradient value higher clustering center, its more for boundary and noise points, usually the number of boundary pixels than noise pixel number, at the same time for boundary point, the $3 \times 3$ neighborhood within close to the pixel gray value of the pixel number is more, its neighborhood similarity is opposite bigger, so you can set its neighborhood characteristic value is 6;And for noise point, within the $3 \times 3$ neighborhood with similar pixel gray value number will be less, the neighborhood similarity is relatively small, so its neighborhood characteristic value is 3.
For regional ant colony, the setting of the heuristic function by the reciprocal of Euclidean distance between pixels of the original to the clustering radius \( r \) and the current search pixels with the ratio of the distance \( d_{ij} \) between clustering center, namely

\[
\eta_{ij} = \frac{r}{d_{ij}} = \frac{r}{\sqrt{\sum_{k=1}^{m} p_k (x_k - \text{center}_k)}},
\]

(2)

By formula can be seen when \( r \) is bigger, the bigger the \( \eta_{ij} \), and clustering, the greater the probability if the current pixel and a clustering center distance, the smaller is \( \eta_{ij} \). \( \eta_{ij} \) actually reflects the current search pixels will be assigned to the expectations of a certain degree. Is the next step in the area of ants routing strategy for:

\[
p_{ij} = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{\sum_{j \in S} [\tau_{ij}]^\alpha [\eta_{ij}]^\beta}
\]

(3)

\( \tau_{ij} \) pixel path pheromone concentration between \( i \) and \( j \).

\( S = \{ X_s | d_{ij} \leq r, s = 1,2,...,N \} \) is the set of all valid path, \( \alpha \) and \( \beta \) area respectively ants accumulated in the process of pixel clustering information expanding rate and heuristic guiding function on the route choice of the impact factor.

For boundary of ant colony, its main task is to regional ant colony search out the compensation for the boundary of the search, to ensure the accuracy of edge detection, boundary ants move depends on differences in the surrounding neighborhood pixels, but in the process of its mobile for boundary identification will still appear error, such as its search pixels, near the border with the real strength of its border similar information could make the ants illusion of real boundary, the need of the ant search border guide, the introduction of the highest two adjacent difference value and the maximum connection similarity factor to select the right path, the heuristic function is as follows

\[
\eta_{ij} = \frac{V(x_i)}{\max \{1, |x_j - x_i|\}}
\]

(4)

\[
V(x_i) = \frac{\sum |x_j - x_i|}{8}
\]

(5)

\( V(x_i) \) here said the difference of adjacent pixels point value, the set of \( NE_j \) said all 8 neighborhood pixels, it can be seen that \( \eta_{ij} \) represents the pixels on the undirected path between \( i \) and \( j \) similar boundary, if \( \eta_{ij} \) represents the existing ideal boundary, to control the search direction of the boundary of ants, if \( V(x_i) = 0 \), is the boundary of the search to the current point ant no longer move.Ants on the route through the left pheromones to communicate with other ants, to guide the action direction of ants, behind the rules referred to as the state transition rules, and attract ants to follow a certain probability is called path selection strategy, the shorter the path will be more and more of the ants pass by, and the legacy of the pheromone concentration will be bigger and bigger, then it will attract more ants on the road, but not all the ants will choose a path that the most attractive, the boundary of ants, get rid of the current information strength biggest
attract to search new path of the path is to be able to search to compensate for boundary conditions, thus, its next path selection strategy for:

$$\text{STR} = \left\{ \arg \max \{ [\tau_{ij}]^\alpha [\eta_{ij}]^\beta \} \right\}$$

$$P_{ij} = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{\sum_{j \in \text{NE}} [\tau_{ij}]^\alpha [\eta_{ij}]^\beta}$$

For type (4) the $\eta_{ij}$ decision boundary path visibility, $\alpha$ and $\beta$, respectively to adjust parameters, $0 \leq q_0 \leq 1$ for maximum impact parameter, it affect the choice of the boundary of ants next direction, if the current compare path probability is less than the maximum relative influence probability, the boundary of ants to search new path, on the other hand, the boundary of ants carried out in accordance with the current path selection probability search next, so the probability of $q_0$ determines the boundary ant to choose the most important boundary of probability, the probability of ants can search to compensation and boundary.

Regional ant colony clustering center according to ants searching for their respective categories produce boundaries, and at the same time to generate its own border assignment of ants, ants in search of compensation with different logo boundary would inevitably encounter other types of boundary ants, at this time if continue to search is likely to search forward to repeat area, this will cause a large number of redundant search, in order to control this kind of situation, adopted the following strategies:

(1) Face to face, belonging to different groups of ants meet two border if the current mobile step will go again to the same pixel, is to stop at this time.

(2) If an ant "touch" to be another ant walking paths, then stop to the direction.

(3) All the ant search to establish a proper tabu table, in the past in which to record their path, so as to find a new path, a predefined time threshold at the same time, to make the ant search into the "dead end" by the time threshold adjustment along the walked the path makes it possible to go back and search again.

(4) The path to the density restriction, near the boundary of ants in the current search path of number greater than a pre-defined value to stop the search.

The specific flow chart is as follows [10-13]
4. The Simulation Experiment and Result Analysis

Experiment parameter is set to $\alpha = 2$, $\beta = 5$, $r=45$, $T = 60$, $\lambda = 0.3$, $\rho = 0.7$, due to the ant to choose the most important probability $q_0$ determines the boundary of probability, so $q_0$ values need to be close to 1, take $q_0 = 0.4$ here, in the boundary of ant colony to compensate boundary search, selection of boundary ant colony number of iterations must weigh the compensation two aspects of the integrity of the boundary and calculation amount, if the number of iterations is insufficient, the incomplete boundary search for compensation, if the number of iterations is more can guarantee the search to complete border, but causes a large amount of calculation [14-15].
Figure 1 (a) as the original image, according to the analysis of the initial clustering center front number can be determined by the histogram to rough, here \( n = 7 \), figure 1 (b) and (c), respectively, using the traditional sobel operator and canny operator to the original image edge detection, and compared with the test results of the worst sobel operator, lost a lot of edge details, canny operator testing is better than sobel operator, but compared with the algorithm proposed in this paper the results still lost a small edge details, figure 1 (d) of the improved ant colony algorithm of edge detection results, can be seen from the comparison results of the improved algorithm are ideal image retrieval effect.
Figure 2 (a) as the original image, the rough number to determine the initial clustering center for n = 10, figure (b) and (c), respectively, for the traditional sobel operator and canny operator test results, figure (d) for the test results of the improved algorithm, the test results can be seen from the diagram sobel operator is still missing many important edge details, but in this section, the proposed algorithm with the test result is the same as the canny operator, but as a result of improved algorithm in the boundary of ant colony set the number of iterations is less, so the efficiency is also faster, in general, the experiment result is satisfactory. Can be seen from the above experimental results and analysis of this chapter puts forward the ant colony algorithm based on region - boundary of ants in the aspect of image retrieval, especially edge detection is better than the traditional test method of performance.

5. Conclusion

Ant colony algorithm is discussed in this paper the application in the field of image retrieval, in view of the traditional ant colony algorithm in edge detection in image retrieval, especially the shortcoming of slow convergence speed and accuracy is not high, on the improvement, proposed one kind of ant colony algorithm based on ant search area - boundary. The algorithm of ant colony was divided into two categories, through
collaboration and mutual cooperation to accomplish effective for image retrieval. Initial in image retrieval, the first ant colony clustering center by using area to preliminary search and retrieval of image, introduced in the late retrieval boundary ant colony through different path selection strategy and strategy of pheromone update to compensate the image boundary search, further improve the detection precision of the algorithm. Two kinds of ant colony with different routing strategies and pheromone update strategy, together to achieve the accurate image retrieval.

References


Author

Hualin Sun, received his Master’s degree in Technology of Computer Application from Southwest Petroleum University in Chengdu, China in 2008. He is currently a lecturer in Changzhou-- Institute of mechatronic technology in Changzhou, China. His research interests include Software component technology, Image processing and algorithm, Cloud Computing, etc.