A Novel Real Time Visual Tracking Method Using Modified Flower Pollination Algorithm

Ong Kok Meng, Ong Pauline, Sia Chee Kiong, Low Ee Soong and Lee Woon Kiow

Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Malaysia.

Abstract. This study presented a novel algorithm, namely, modified flower pollination algorithm (MFPA), for moving object tracking in real time video. Three variables, i.e., x-coordinate, y-coordinate and width, known as search window, were used to represent the features of the target object. The image within the search window was extracted, and used to calculate the Hue, Saturation, and Value (HSV) histogram. The HSV histogram was then used as a feature model of the target object. Subsequently, the MFPA will search for an optimal match of the target using Bhattacharyya distance as the fitness function. Simulation results demonstrated that the MFPA outperformed FPA in the aspect of higher accuracy, where 75.35% and 92.13% of improvements were observed in two considered case studies.

1. Introduction
Due to high-powered computers proliferation, the existing of low cost and high quality video cameras as well as the high demand for automated video analysis have made object tracking a crucial subject in computer vision (Yilmaz, Javed, & Shah, 2006). Furthermore, object tracking is being found useful in real-world applications, such as visual surveillance, human computer interaction, medical imaging, video compression, service robot and traffic flow monitoring (Walia & Kapoor, 2014). However, dealing with the changes of complex object appearance in real time video such as illumination variation, partial occlusion, shape deformation and camera motion, still remained an open issue (Li et al., 2013).

In object tracking, the target is selected in the first frame as input and modelled mathematically so that the visual tracking algorithm can make decision to determine the position of the target in next frame (Gao, Yin, Zou, Li, & Liu, 2015). This process is considered as appearance modelling. In next frame, the visual tracking algorithm will find the position of the target which has the most similar feature as the target in first frame, and this process is considered as searching strategy. Therefore, appearance modelling and searching strategy are important in tracking the target precisely in continuous sequence of images from the video (Li et al., 2013). Hence, a good tracker with an accurate appearance modelling and an efficient searching strategy can efficiently locate the most similar candidate region of the target.

There are two modules under appearance modelling, namely, the visual representation and statistical modelling. Visual representation focuses on how to define the characteristics of the target object appearance, while statistical modelling concentrates on how to build a robust statistical model using different types of statistical learning schemes (Li et al., 2013).

Besides appearance modelling, researchers have focused on searching strategy. In tracking process, target is being searched in every possible candidate region within the image. With the help of efficient searching strategy, target can be positioned precisely in an image within short period of time.
Searching strategy can be divided into two methods - probabilistic method and deterministic method (Gao, He, Luo, & Yu, 2012). Probabilistic method views the tracking algorithm as a state solving problem under a Bayesian framework, modeling uncertainty and propagating the conditional densities through the tracking process (Gao et al., 2012). Examples of probabilistic method are Kalman Filter (Kalman, 1960) and Particle Filter (Arulampalam, Maskell, Gordon, & Clapp, 2002). Deterministic method searches the target in each frame iteratively by maximizing the similarity measure between this region and the target model as well as estimating the position of target in the following frame (Gao et al., 2012). Examples of deterministic method are Mean Shift (Dorin Comaniciu, Ramesh, & Meer, 2003) and Trust Region (Liu, 2004).

It is obviously that visual tracking process can be considered as one of the optimization problems (Fourie, Mills, & Green, 2010). The fitness function of the visual tracking process can be defined by comparing the searching candidate and the target model. Therefore, researchers start to apply metaheuristic optimization algorithm as searching strategy, to replace those traditional tracking strategies in visual tracking process (Bae, Kang, Liu, & Chung, 2016; Gao et al., 2016; Sardari & Moghaddam, 2016). This is due to the advantages of these algorithms can track the target efficiently without making any assumptions during the search process, such as the shape of distribution or the noise in the system (Gao et al., 2012).

In recent years, flower pollination algorithm (FPA) (Yang, 2012) has been widely used in optimizing real world problems, such as minimizing the weight of truss structures and searching the optimum mass (Bekdas, Nigdeli, & Yang, 2015), optimizing the cold extrusion process (Ong, Vui, Ho, & Ng, 2016), and optimizing period and damping ratio of tuned mass dampers for designing the seismic structures (Nigdeli, Bekdas, & Yang, 2016). Results showed that FPA can converge faster toward the optimal solution than other algorithms. In addition, preliminary results in pioneer work of FPA demonstrated that FPA outperformed genetic algorithm (GA) and particle swarm optimization (PSO) in solving ten benchmark test functions. Therefore, in this study, FPA is applied in visual tracking process.

FPA, however, has its own limitations. The first limitation is the random generated initial population in FPA. Second limitation is the randomization of the local pollination process. The last limitation is the random walk of the global searching of Lévy flight. In order to tackle these limitations in this study, chaos theory with circle map has been used to initialize the initial population in FPA (Ong, Ong, & Sia, 2016a). Second limitation associated to local searching is overcome with the integration of frog leaping local search. Finally, inertia weight is integrated to Lévy flight to solve the third limitation. In this regard, the resultant algorithm, named as modified flower pollination algorithm (MFPA), will be used for visual tracking process.

This paper is organized as follows: Section 2 describes the details of the proposed MFPA. In Section 3, the implementation of MFPA in visual tracking system is explained. The experimental results are presented in Section 4 and lastly, Section 5 is devoted to conclusion.

2. Modified Flower Pollination Algorithm

The improvement strategies in MFPA are three-folds. Firstly, the initial population is generated using circle map, in order to improve the diversity of initial population. Secondly, the local search of FPA is replaced with frog leaping local search. Lastly, the inertia weight is integrated to Lévy flight in global pollination (Ong, Ong, & Sia, 2016b). The flow involved in MFPA is described as follows.

Prior to the searching process, the parameters of population size, \( n \), dimension of search space, \( d \), maximum iteration, \( \text{max}_\text{iter} \), switch probability \( p \), range of search space, \([Lb, Ub]\), number of memeplexes, \( m \) and iterations within each memeplex, \( it \), are initialized. Subsequently, circle map is used to initialize MFPA initial population, with

\[
x_{n+1} = (x_n + 0.2 - (0.5/2\pi)\sin(2\pi x_n)) \mod(1)
\]

\[
X_{ij} = Lb + (Ub - Lb)x_{ij}
\]
where \( x_n \) is the generated solution by circle map, while \( X_{ij} \) is the remapped initial population based on lower bound and upper bound of design variables. The best solution \( g_1 \) is then identified and its fitness value is evaluated.

Next, the entire population will undergo the frog leaping local search. Shuffled Frog-Leaping Algorithm (SFLA) is developed by Eusuff, Lansey & Pasha, (2006), inspired from the foraging behaviors of frogs (Eusuff et al., 2006). The SFLA begins with initializing a population of frog in a wetland, which is then divided into several small groups. Information is shared between the best frog and the worst frog within the group. When each group has independently searched for food over a period of time, all groups are gathered and formed a population. This process is repeated until termination condition is achieved.

The location of the worst frog is updated using:

\[
\text{Change in frog position (D)} = \text{rand} \cdot C \cdot (X_b - X_w)
\]

\[
\text{New position } X_w = \text{current position } X_w + D_i, \quad (D_{\text{max}} \geq D_i \geq -D_{\text{max}})
\]

where \( \text{rand} \) is a random number chosen between 0 and 1, \( C \) is a search-acceleration factor, \( X_b \) is best solution, \( X_w \) is worst solution and \( D_{\text{max}} \) is the maximum allowed change in a frog’s position.

Following from the local search using SFLA, if a generated random number is greater than the switching probability \( p \) (\( \text{rand}>p \)), the global search using

\[
x_{i+1} = x_i + wL(x_i - g_*)
\]

takes place. \( L \) is Lévy distribution while \( w \) is the inertia weight determined from

\[
w = w_{\text{min}}(1 + \frac{w_{\text{max}}}{\sqrt{t}} \cdot \text{tanh}(\frac{g^*}{g_i})))
\]

where \( t \) denotes the current iteration number and \( g^* \) is the current fitness value. If \( \text{rand} \leq p \), the solution will remain in its position. The introduction of inertia weight into Lévy flight can expand the MFPA search space at the early phase so that it will not easily get trapped in local minima and increase the convergence rate at the later phase due to the value of \( w \) will be decreased as the iteration increases.
Subsequently, the fitness value of each new solution is evaluated. The current solution is replaced with the new solution if the fitness value of the new solution is better, or else, the solution remains unchanged. The similar updating is applied to the global best solution as well. The process is repeated until the termination condition is satisfied (Ong, Ong, et al., 2016b). The flowchart of MFPA is shown in Figure 1.

3. Visual Tracking using Modified Flower Pollination Algorithm
To implement MFPA in visual object tracking, first, the target in a video is tracked by using a rectangle around it, defined by three dimension variables (x-coordinate, y-coordinate and width). Since the ratio between the height and width of the target is constant throughout the whole video, the height-width ratio is identified in the beginning of the visual tracking system. Then, the appearance modelling of the target object is established using colour histogram.

As the motion of the target object is continuous in most cases, the target object is assumed to be moving slightly to a place near its previous position between two consecutive frames. In order to enhance the performance of the MFPA, a reason-able size of the sub-search area is given as lower boundary, Lb and upper boundary, Ub to those three dimension variables in MFPA parameter initialization. The equations are:

\[ L_{\text{b,x}} = x_{\text{frame,i}} - T_w \]  
\[ L_{\text{b,y}} = y_{\text{frame,i}} - T_h \]  
\[ L_{\text{b,w}} = \text{width}_{\text{frame,i}}(1 - \alpha) \]  
\[ U_{\text{b,x}} = x_{\text{frame,i}} + T_w \]  
\[ U_{\text{b,y}} = y_{\text{frame,i}} + T_h \]  
\[ U_{\text{b,w}} = \text{width}_{\text{frame,i}}(1 + \alpha) \]

where \( T_w \) and \( T_h \) are predefined parameters that regulate the size of the search-space and \( \alpha \) is another predefined parameter that controls the size of width for current solution.

Subsequently, the MFPA as explained in Section 2 is used to track the target in current frame. Fitness function is one of the important components in optimization problem. The fitness function of visual tracking system is to measure the similarity between the solution in current frame and the target object in previous frame. In this study, colour histogram is used as the criteria to evaluate the similarity between the solution and the target object. Specifically, Bhattacharyya distance is used to measure the similarity of those two histograms, given by

\[ \text{Fitness Function} = \sqrt{1 - \frac{\sum_{n} \sqrt{H_i(n)H_t(n)}}{\sqrt{H_i H_t N^2}}} \]

where \( N \) is the total number of histogram bins, \( H_i \) is the histogram of the solution \( i \), \( H_t \) is the histogram of the target object, \( n \) represents the distribution domain which the range of value is from 0 to 255. \( H_i(n) \) and \( H_t(n) \) are pixel numbers with a specific HSV value \( n \) for the solution and target object, respectively (D Comaniciu, Ramesh, & Meer, 2000).

The obtained global best solution is regarded as the tracked target, and its x-coordinate, y-coordinate and width are used to update the parameters in Equations (7) to (12). The process is repeated until the end of video sequences.
3.1. Experimental Setup

MFPA and FPA are implemented in visual tracking system and their results are evaluated and compared. The same video and fitness function are used to evaluate both algorithms. Additionally, if the global best solution in previous frame is less than 0.03, the searching area of current frame is 15% of the size-x and size-y of the frame. This is due to the target object is considered nearby when the fitness value is less than 0.03. Hence, reducing the search area can save the computational cost. Also, when the value of the global best solution in the previous frame is increasing due to the occlusion, the search area of current frame, the number of initial population and the number of iteration required by the algorithm will be increased, to enable the algorithm to find the target object precisely.

In this study, two videos, namely, biker and dog from Visual Tracker Benchmark (http://cvlab.hanyang.ac.kr/tracker_benchmark/datasets.html) with resolution of 640 x 360 and 352 x 240, respectively, are used to test the performance of MFPA and FPA. The computer used for object tracking has an Intel ES-1220VS 3.5GHz Xeon processor with 32GB RAM. The target objects of biker and dog video selected at the first frame of the video are shown in Table 1.

Table 1 shows a biker with green colour shirt and green colour background on the railway. The biker cycles from far toward the camera on the left hand side of the railway and shifts his path toward right hand side of the railway and moves away from the camera. In frame 50 and 71, MFPA has tracked the target precisely, as compared to FPA. Furthermore, in frame 78, 85 and 106, FPA has lost the target object but MFPA still could track the target object accurately. Therefore, the performance of MFPA is significantly better due to its successive rate in object tracking, without influenced by background colour that is similar to the target.

Table 1 also shows a dog running toward its owner. In frame 12 and 28, it can be observed that MFPA can track the dog better than the FPA. In frame 50, 57 and 64, MFPA has correctly tracked the position of the dog meanwhile FPA has lost its target. In frame 85 and 124, the dog is still located within the red rectangular box using MFPA but FPA fails to do so. In short, MFPA has shown a better object tracking ability with less number of function evaluation as compared to the FPA.

3.2. Position Errors

In this part, the correctness of the tracked object x-position and y-position is compared between FPA and MFPA. The results are shown in Figure 2 to Figure 5. In Figure 2 and Figure 3, it can be observed that the performance of FPA is unstable starting from frame 70 to 140 as the line produced by FPA keeps fluctuating along those frames. Meanwhile, the performance of MFPA is stable starting from beginning until the end, as the line produced by MFPA is almost the same as the actual position line.

In Figure 4, the line produced by FPA keeps oscillating from frame 40 to 70 whereas the line produced by MFPA follows the actual position line (blue). In Figure 5, it is obvious that the line produced by FPA has small oscillation from beginning until the end whereas the line produced by MFPA follows the actual position line closely. Thus, MFPA has shown its higher accuracy as well as stability by observing Figure 2 to Figure 5.

3.3. Fitness Value

In Figure 6, the fitness value obtained by the MFPA is lower than FPA in 75.35% from the total frames. In Figure 7, the fitness value obtained by MFPA in 92.13% from the total frames is lower than the FPA. Therefore, in general point of view, MFPA has performed better than FPA in visual tracking system.
Table 1. Biker and Dog Tracking Results From FPA and MFPA

| Frame Index | Biker Tracking Result | Frame Index | Dog Tracking Result |
|-------------|-----------------------|-------------|--------------------|
|             | FPA       | MFPA       | FPA            | MFPA            |
| 1           | ![Image]  | ![Image]  | 1              | ![Image]  |
| 30          | ![Image]  | ![Image]  | 12             | ![Image]  |
| 50          | ![Image]  | ![Image]  | 28             | ![Image]  |
| 71          | ![Image]  | ![Image]  | 50             | ![Image]  |
| 78          | ![Image]  | ![Image]  | 57             | ![Image]  |
| 85          | ![Image]  | ![Image]  | 64             | ![Image]  |
| 106         | ![Image]  | ![Image]  | 85             | ![Image]  |
| 133         | ![Image]  | ![Image]  | 124            | ![Image]  |

Average Number of Function Evaluation

| FPA | MFPA |
|-----|------|
| 51  | 41   |

| FPA | MFPA |
|-----|------|
| 51  | 41   |
**Figure 2.** X-position errors (pixel) of video – Biker

**Graph of X-Coordinate Against Frame**

- Actual x
- MFPA x
- FPA x

Frame 0 to 150

**Figure 3.** Y-position errors (pixel) of video – Biker

**Graph of Y-Coordinate Against Frame**

- Actual y
- MFPA y
- FPA y

Frame 0 to 150
Figure 4. X-position errors (pixel) of video – Dog

Figure 5. Y-position errors (pixel) of video – Dog
4. Conclusion

In this study, a novel MFPA is used to track a moving target object in video sequences. In MFPA, the solutions evolve their position based on the frog leaping local search and improved global pollination. Due to the information sharing among frogs in frog leaping local search, the convergence rate of MFPA is higher than FPA. In addition, the improved global pollination has improved and controlled the search space exploration more entirely. Besides that, the initial population is generated using circle map, providing non-repetition and ergodicity solutions has further enhanced its performance. In this regard, MFPA has the ability of tracking the target object precisely that is 75.35% and 92.13% better
than FPA in bike and dog case study, respectively, with less number of function evaluation as compared to FPA which is 19.61% less from the number of function evaluation of FPA as shown in Table 1, Figure 6 and Figure 7.

5. References

[1] Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2), 174–188. https://doi.org/10.1109/78.978374

[2] Bae, C., Kang, K., Liu, G., & Chung, Y. Y. (2016). A novel real time video tracking framework using adaptive discrete swarm optimization. *Expert Systems with Applications*, 64, 385–399. https://doi.org/10.1016/j.eswa.2016.08.027

[3] Bekdas, G., Nigdeli, S. M., & Yang, X. (2015). Sizing optimization of truss structures using flower pollination algorithm.

[4] Comaniciu, D., Ramesh, V., & Meer, P. (2000). Real-time tracking of non-rigid objects using mean shift. *IEEE Conference on Computer Vision and Pattern Recognition*, 2(7), 142–149. https://doi.org/10.1109/CVPR.2000.854761

[5] Comaniciu, D., Ramesh, V., & Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5), 564–577. https://doi.org/10.1109/TPAMI.2003.119591

[6] Eusuff, M., Lansey, K., & Pasha, F. (2006). Shuffled frog-leaping algorithm: A memetic metaheuristic for discrete optimization. *Engineering Optimization*, 38(2), 129–154. https://doi.org/10.1080/03052150500384759

[7] Fourie, J., Mills, S., & Green, R. (2010). Harmony filter: A robust visual tracking system using the improved harmony search algorithm. *Image and Vision Computing*, 28(12), 1702–1716. https://doi.org/10.1016/j.imavis.2010.05.006

[8] Gao, M.-L., He, X.-H., Luo, D.-S., & Yu, Y.-M. (2012). Object tracking based on harmony search: comparative study. *Journal of Electronic Imaging*, 21(4), 043001–1. https://doi.org/10.1117/1.4714301

[9] Gao, M.-L., Shen, J., Yin, L.-J., Liu, W., Zou, G.-F., Li, H.-T., & Fu, G.-X. (2016). A novel visual tracking method using bat algorithm. *Neurocomputing*, 177, 612–619. https://doi.org/10.1016/j.neucom.2015.11.072

[10] Gao, M.-L., Yin, L.-J., Zou, G.-F., Li, H.-T., & Liu, W. (2015). Visual tracking method based on cuckoo search algorithm. *Optical Engineering*, 54(7), 073105. https://doi.org/10.1117/1.OE.54.7.073105

[11] Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering, 82*(1), 35. https://doi.org/10.1115/1.3662552

[12] Li, X., Hu, W., Shen, C., Zhang, Z., Dick, A., & Hengel, A. van den. (2013). A Survey of Appearance Models in Visual Object Tracking, 1–42. https://doi.org/10.1145/2508037.2508039

[13] Liu, T. (2004). Real-Time Tracking Using Trust-Region Methods, 26(3), 397–402.

[14] Nigdeli, S. M., Bekdas, G., & Yang, X.-S. (2016). Optimum tuning of mass dampers for seismic structures using flower pollination algorithm, I, 264–268.

[15] Ong, P., Ong, K. M., & Sia, C. K. (2016a). An Improved Flower Pollination Algorithm with Chaos Theory for Function Optimization. Batu Pahat, Johor.

[16] Ong, P., Ong, K. M., & Sia, C. K. (2016b). Application of Modified Flower Pollination Algorithm on Mechanical Engineering Design Problem. Batu Pahat, Johor.

[17] Ong, P., Vui, D., Ho, C. S., & Ng, C. H. (2016). Modeling and optimization of cold extrusion process by using response surface methodology and metaheuristic approaches. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-016-2626-8

[18] Sardari, F., & Moghaddam, M. E. (2016). An object tracking method using modified galaxy-based search algorithm. *Swarm and Evolutionary Computation*, 30, 27–38. https://doi.org/10.1016/j.swevo.2016.04.001
[19] Walia, G. S., & Kapoor, R. (2014). Intelligent video target tracking using an evolutionary particle filter based upon improved cuckoo search. *Expert Systems with Applications, 41*(14), 6315–6326. https://doi.org/10.1016/j.eswa.2014.03.012

[20] Yang, X.-S. (2012). Flower Pollination Algorithm for Global Optimization.

[21] Yilmaz, A., Javed, O., & Shah, M. (2006). Object Tracking: A Survey. *ACM Computing Surveys, 38*(4), 13–es. https://doi.org/10.1145/1177352.1177355

**Acknowledgments**

Financial support from Universiti Tun Hussein Onn Malaysia (UTHM) in the form of IGSP Grant Vot U671 and GPPS Vot U806 are gratefully acknowledged.