

Particle Swarm Optimization in Foreground Segmentation for Live Video

Harsha M.¹, Mithra S T²

¹M.Tech student, Department of Communication Engineering, Sree Buddha College of Engineering for Women, Elavumthitta, Kerala, India

²Assistant Professor, Department of Electronics and Communication Engineering, Sree Buddha College of Engineering for Women, Elavumthitta, Kerala, India

Abstract: *The goal of foreground segmentation is to extract the desired foreground object from input videos and matting problem means the problem of extracting the foreground object accurately. Over the years, there have been significant amount of efforts on how to extract objects from live videos. The support vector machine (SVM) classification is an active research area which solves classification by finding the best hyperplane that separates all data points of one class from those of the other. In many problems of classification, the performances of SVM are often evaluated by the rate of error depends on the optimization method adopted to label the unknown pixels along the boundary. Here, among the methods of different optimization methods, selected the method of PSO (Particle Swarm Optimization) which makes it possible to optimize the performance of classifier and it will enhance the accuracy of the foreground object segmentation of live videos. PSO is a population based stochastic search process, shaped after the social behavior of a bird flock. It is similar to flock of birds migrating towards some destination, where the intelligence and quality lies in the co-operation of an entire flock. So over the iteration, a group of variables have their values adjusted closer to member whose value is closest to the target at any given moment. Foreground segmentation is addressed in the application of new background substitution and shown to create convincingly high quality composite video.*

Keywords: Foreground segmentation, Matting, Classifier, Support vector machine, Particle swarm optimization.

1. Introduction

Video segmentation, studies how to extract objects of interest from input videos in visual signal processing research. The goal of foreground segmentation is to simplify and change the representation of the video into something that is more purposeful and easier to analyze. Video segmentation is typically used to locate objects and boundaries in videos. It is a fundamental problem in real-time applications and often performs as a pre-analyzing step for other video analysis tasks such as background substitution, teleconferencing, action recognition and retrieval. Over the years a significant amount of related techniques have been proposed in both computer vision and graphics communities. However, some of them are restricted to sequences captured by stationary cameras, and others require significant amount of training examples or cumbersome user interactions. As a result, there still lacks an efficient and competent algorithm capable of processing demanding live video scenes with minimum user interactions. Motivated by the above finding here present a novel method for foreground segmentation and boundary matting. The key insight of the approach is to maintain support vector machines at every pixel location which capture the local foreground and background color densities separately and jointly label pixel either as a background or foreground [1].

In many issues of classification, the performances of SVM are often evaluated by the rate of error depends on the optimization method adopted. To address this issue, among the methods of optimization methods chose particle swarm optimization and it will enhance the accuracy of the foreground object segmentation for live video. Accuracy of

segmentation is derived in terms of sensitivity and specificity. The ability to correctly detect pixel with true foreground and background is refers to sensitivity. Otherwise it is refers to specificity. PSO is a population based stochastic search process, shaped after the social behavior of a bird flock. It is similar in spirit to flock of birds migrating towards some destination, where the intelligence and quality lies in the co-operation of an entire flock. PSO is used to find optimal feature subsets by discovering the best feature combinations within the problem space by following the particles with the best solutions so far. So over the iteration, a group of variables have their values adjusted closer to member whose value is closest to the target at any given moment [2]. Foreground segmentation is addressed in the application of new background substitution and shown to generate convincingly high good quality composite video.

2. Proposed System

The proposed system aimed on real-time foreground segmentation and boundary matting for live videos with SVM technique using PSO. Support vector machine is the supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification of pixels. PSO is used to find optimal feature subsets by discovering the best feature combinations within the problem space by following the particles with the best solutions so far. Like any evolutionary algorithm, the algorithm of optimization is influenced by factors such as the criterion of stop, the objective function. Iteration count attached to the precondition can be considered as the criterion of stop. The purpose of the objective function will be to reduce the error of generalization to the minimum.

Figure 1 shows the architecture of foreground segmentation for video.

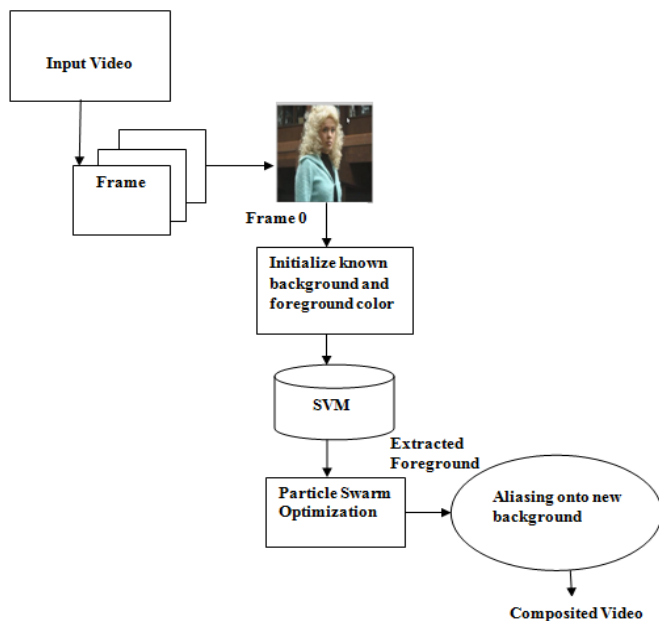


Figure 1: Architecture of foreground segmentation of video

2.1 Support Vector Machine

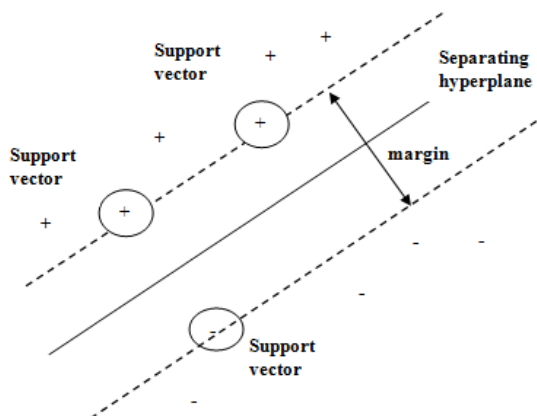


Figure 2: Optimal separating hyperplane

A simple representation of SVM is shown in Figure 2. The linear classifier is termed the optimal separating hyperplane. The support vector machine (SVM) classification is an active research area which solves classification by finding the best hyperplane that separates all data pixels of one class from those of the other. The linear hyper plane is used to create a classifier with a maximum margin. That is one for foreground and the other for background. The SVM is trained locally for each pixel using known foreground and background colors. Once trained, it is used to jointly label pixel either as foreground, background, or unknown. The best hyperplane for an SVM means the one with largest margin between the two classes (background and foreground). Margin means the maximal width of the slab parallel to the hyperplane that has no interior data points. The support vectors are the data points that are closest to the separating hyperplane between the two pixel classes.

2.2 Particle Swarm Optimization

There may be pixels in the frame with colors that are not recognized by SVM. These pixels are labeled as unknown at the end of the process. Our next task is to label these unknown pixels along the boundary by optimization so that a clean binary segmentation can be generated.

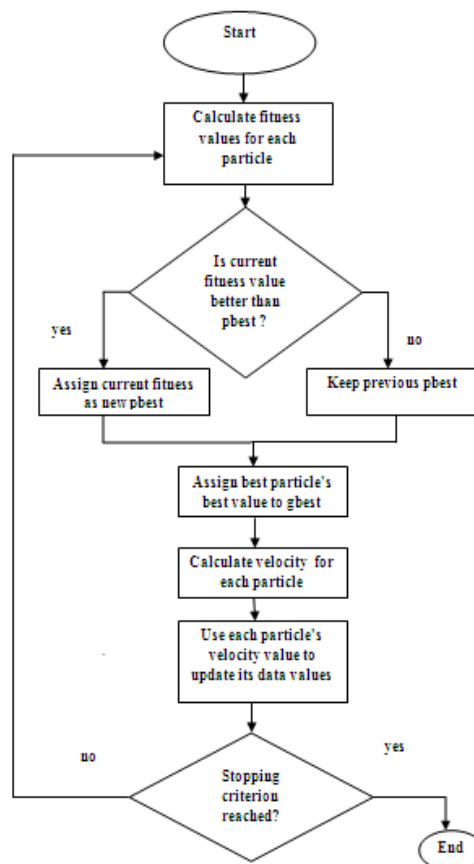


Figure 3: Flow diagram illustrating the particle swarm optimization algorithm.

Figure 3 shows the flow diagram of PSO. Inspired by the flocking and schooling patterns of birds and fish, particle swarm optimization was invented by Russell Eberhart and James Kennedy. Initially, these two started out developing computer software simulations of birds flocking around food sources, and then later realized how well their algorithms worked on optimization problems. PSO might sound complicated, but it's really a very simple algorithm. PSO is a population based stochastic search process, shaped after the social behavior of a bird flock. It is similar to flock of birds migrating towards some destination, where the intelligence and quality lies in the co-operation of an entire flock. So over the iteration, a group of variables have their values adjusted closer to member whose value is closest to the target at any given moment. It's an algorithm that's simple and easy to implement. The velocity is calculated according to how far a particle's data is from the target.

The individuals furthest from the target would make an effort to keep up with the others by flying faster toward the global best particle. The velocity would mark out how different the pattern is from the target, and thus, how much it needs to be adapted to match the target. Each individual's personal best

(pbest) value only designates the closest the data has ever come to the target since the algorithm started. The global best (gbest) value only modifies when any particle's pbest value comes closer to the target than gbest. As iterations progresses, gbest gradually moves closer and closer to the target until one of the particles reaches the target [2].

Through each iteration, the algorithm keeps track of three global variables: target value or condition, global best and the stopping value. Global best value tracing which particle's data is currently closest to the target and the stopping value marking when the algorithm should stop if the target isn't found. Each particle consists of a data depicting a feasible solution, a velocity value indicating how much the data can be changed, and a pbest value indicating the near the particle's data has ever come to the target.

The velocity and position update step is decision making for the optimization ability of the PSO algorithm. The velocity of each individual in the swarm is updated using the following equation:

$$v_i(t+1) = wv_i(t) + c_1r_1\left[\hat{x}_i(t) - x_i(t)\right] + c_2r_2[g(t) - x_i(t)] \quad (1)$$

The particle's index is represented by i . Thus, velocity of particle (say i) is $v_i(t)$ at time t and the position of particle is represented by $x_i(t)$ at time t . The parameters w , c_1 , and c_2 ($0 \leq w \leq 1.2$, $0 \leq c_1 \leq 2$, and $0 \leq c_2 \leq 2$) are user-supplied coefficients. The values r_1 and r_2 ($0 \leq r_1 \leq 1$, and $0 \leq r_2 \leq 1$) are random values regenerated for each velocity update. The individual best candidate solution for particle i at time t is represented by $\hat{x}_i(t)$, and the swarm's global best candidate solution at time t is $g(t)$. r_1 in the cognitive component and r_2 in the social component are the random values cause these components to have a stochastic influence on the velocity update. This stochastic optimization nature causes each individual to move in a semi-random manner heavily influenced in the directions of the individual best solution of the particle and global best solution of the swarm. The first, second and third term are known as inertia component, cognitive component and social component respectively [3].

Once the velocity for each particle is computed, each particle's position is updated by applying the new velocity to the particle's previous position:

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (2)$$

This process is repeated until some stopping condition is reached. Some common stopping conditions include: a number of iterations since the ultimate update of the gbest candidate solution, a preset number of iterations of the algorithm, or a predefined target fitness value. In order to keep the particles from moving too far beyond the search space used a method called velocity clamping to limit the maximum velocity of each particle.

3. Simulation Results

Matlab R2013a is used as the platform to perform this task.



Figure 4: Input Video



Figure 5: Draw a mask



Figure 6: Segmented Video



Figure 7: Composited Video

Figure 4 shows the input video to be processed for the foreground segmentation. Firstly read all the video frames and get the number of frames. Then create a user defined mask on the video as shown in Figure 5. As time goes the mask will grow outwards. When it reaches the boundary of the foreground object, SVM sets outside of the mask (background) as zero and inside of the mask (foreground) as such. The foreground object segmented for live video is shown in Figure 6. Finally the composite video is shown in Figure 7. Set $w = 1.2$, $c_1 = c_2 = 0.8$, number of iteration = 150 and PSO population = 150.

Table 1 shows the performance of the segmentation in terms of accuracy, sensitivity and specificity. Sensitivity, specificity and accuracy are described in terms of true positive (TP), true negative (TN), false negative (FN) and false positive (FP). A good binary classification test always results with high values for all the three factors [4].

Table 1: Performance measures

Accuracy	Specificity	Sensitivity
98.741	97.044	99.00

4. Conclusion

A system and method imposed as a software tool for foreground segmentation of video sequences in real-time and SVM used as a classifier which separates pixels into two categories (foreground and background). In many problems of classification, the performances of SVM are often evaluated by the rate of error depends on the optimization method adopted. To address this issue, PSO is used to resolved unknown or boundary conditions in order to achieve accurate segmentation. PSO is used to find optimal feature

subsets by discovering the best feature combinations as they fly within the problem space by following the particles with the best solutions so far. Accuracy of segmentation is derived from sensitivity and specificity. As a conclusion, the results showed the accuracy has increased, as well as sensitivity and specificity of the segmentation.

5. Acknowledgment

I would like to express profound gratitude to our Head of the Department, Prof. Cherian Schariah, for his encouragement and for providing all facilities for my work. I express my highest regard and sincere thanks to my guide, Asst. Prof. Ms. Mithra S T, who provided the necessary guidance and serious advice for my work.

References

- [1] Minglun Gong, Yiming Qian and Li Cheng, "Integrated Foreground Segmentation and Boundary Matting for Live Video," IEEE Transactions on Image Processing, vol.24, no.4, pp. 1356-1370, April 2015.
- [2] Fatima Ardjani, Kaddour Sadouni, "Optimization of SVM Multiclass by Particle Swarm (PSO-SVM), " I.J.Modern Education and Computer Science Press, University of Sciences and Technology, 2010.
- [3] Soodabeh Darzi, Tiong Sieh Kiong and Balasem Salem, "Overview of Particle Swarm Optimization on its Applications and Methods, " Aust. J. Basic & Appl. Science, pp. 490-499, 2013.
- [4] Mokhairi Makhtar, "Optimizing Sensitivity And Specificity Of Ensemble Classifiers For Diabetic Patients," JATIT, vol.82. No.2, pp. 230-236, 2015 .
- [5] Cheng and M. Gong, "Realtime background subtraction from dynamic scenes," in Proceedings of the IEEE on Comput. Vis., pp. 2066–2073, 1999.
- [6] Y. Sheikh and M. Shah, " Bayesian object detection in dynamic scenes, " in Proceeding of IEEE on Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 74–79, 2005.
- [7] Gong and L. Cheng, "Real-time foreground segmentation on GPUs using local online learning and global graph cut optimization," in Proceedings of Int. Conf. Pattern Recognit., pp. 1–4, 2008.
- [8] J. Zhong and S. Sclaroff, "Segmenting foreground objects from a dynamic textured background via a robust Kalman filter," in Proceeding of. 9th Int. Conf. Comput. Vis., pp. 44–50, 2003.
- [9] J. Wang, P. Bhat, R. A. Colburn, M. Agrawala, and M. F. Cohen, "Interactive video cutout," ACM Trans. Graph., vol. 24, no. 3, pp. 585–594, 2005.
- [10] J.Wang and M. F. Cohen, "An iterative optimization approach for unified image segmentation and matting," in Proceeding of 10th IEEE Int. Conf. Comput. Vis., pp. 936–943, 2005.
- [11] www.mathworks.in

Author Profile

Harsha M. received the B-Tech degrees in Electronics and Communication Engineering from M.G University, Kerala at Sree Buddha College of Engineering for Women in 2014. And now she is pursuing her M-Tech degree in Communication Engineering under the same university in Sree Buddha college of Engineering for women.

Mithra S T working as Assistant Professor in department of Electronics and Communication, Sree Buddha College of Engineering for Women, Elavumthitta, Pathanamthitta.