2021-1

by Muniri Muniri

PAPER · OPEN ACCESS

The improvement of visual object tracking using a combination between particle and correlation filter

View the article online for updates and enhancements.

The improvement of visual object tracking using a combination between particle and correlation filter

Yabunayya Habibi and Muniri

Mathematics Education Department, State Islamic Institute of Tulungagung, Tulungagung, Indonesia

E-mail: habibiexebppt@gmail.com, muniri@iain-tulungagung.ac.id

Abstract.   One of the many challenging problems from computer vision is constructing robust object tracking from various problems and object conditions. Difficulties in tracking are more complex when an image sequence has many types of tracking problems such as background clutter, fast motion, occlusion, illumination variance, scale variance, motion blur, non-rigid object pattern, and another. Meanwhile, the task to build a novel approach which efficiently handle all tracking problems still in development step, especially supported by low computing time. We combine adaptive particle filters (APF) and discriminative correlation filters (DCF) to overcome more complex tracking problems so it can improve overall tracking accuracy. Our combination employs APF as the main framework and embeds DCF in it to estimate motion vector. However, this will make a double-edged knife where the utilization of two approaches together will make enlarged computing time. So, we also utilize leap-tracking as an alternative solution to reduce high computation time. Our testing uses 15 sets of benchmarks tracking datasets and 3 sets by ourselves with various tracking problems. The results show that our combination significantly improves tracking accuracy. Thus, the reducing of computational procedures efficiently will increases tracker speed without reduces accuracy and avoids identification errors.

1. Introduction

Robust object tracking is a challenge in the computer vision field, especially in its many applications such as robotics, navigation supervision, human-computer interaction and automation. Generally, object tracking runs based on initialization described earlier. Then, the task is estimating object motion in a image sequence with a hope where each object in an image can be tracked precisely. Even though some studies have proposed tracking approach for last time but there are still many things which need to be developed to overcome various tracking problems such as fast motion, occlusion, illumination variance, scale variance, motion blur, non-rigid object pattern and another [1].

There are many approaches or designs for object tracking to overcome object tracking problems. Swalaganata et al [2] and Habibi et al [3] construct a tracking on a small object target. In that approach, tracker is assisted by super-resolution to obtain more information from the object. Swalaganata et al [2] use an HSV-based camsift which utilizes a histogram with varying conditions and colors. Meanwhile, Habibi et al [3] use the Adaptive Particle Filter which utilizes an adaptive motion model. However, both use different super-resolution techniques, single-frame and multi-frame. Both approaches obtain a significant increase in precision.

Recently, the development of correlation filters for tracking objects is very fast and evidence shows very good results [4][5] [6]. Moreover, correlation filters have the advantage of accuracy for visual problems such as object rotation, scale variance and efficiency in terms of computational time, thereby their application makes it very suitable for a variety of real-time applications. Tracking by a correlation filter like DCF, exploits FFT to get an optimal correlation filter. An adaptive correlation filter from Bolme et al [7] based on the output the minimum Sum of Squared Error (MOSSE) compiled via convolution. Based on MOSSE, Danelljan et al [4] construct separate learning between translation and scale for computational time efficiency.

Meanwhile, APF also has many variations because of its development in overcoming the problem of tracking. In this paper, we use APF by [8], which adapt an adaptive motion model where a number of particles with its weight as a representation of a candidate position support object to obtain a good particle distribution. A filter considers the weight of particles with visual disturbance, motion continuity, trajectory smoothness, and combines them in a likelihood of observation. Thus, that conditions support to overcome visual disturbances such as background clutter, motion blur, and fast object motion. In addition, a prediction in APF is effective in overcoming occlusion and avoiding motion estimation errors.

Although the correlation filter achieves competitive results in accuracy and computational time efficiency, there are still some tracking problems that cannot be overcome by DCF. In this paper, we propose a combination of APF and DCF to solve more complex tracking problems. In addition, an alternative solution is also proposed to reduce the high computation time as a consequence of the combination of the two approah.

2. Tracking system

This section will review the literature shortly about the tracking approach that we use. Section 2.1 reviews the APF by [8] how this tracking works as the main framework of tracking, especially in visual disturbance conditions, motion continuity of the objects, trajectory smoothness, and predicting the object position in occlusion or object motion which aren't detected. Then, section 2.2 reviews the DCF by [4]

2.1 Adaptive particle filter

A particle filter is a state space which implements a recursive Bayes filter with posterior probability. Posterior probabilities are estimated using weighted particle sets. Each particle represents the center candidate object $X_t = (x, y)$ which corresponds to the weight of the particle. Determination of X_t based on weighted average of all particles. The reliability of the particle filter depends on the effectiveness of the observation model of the particle weight.

A moving object is modeled as

$$X_{t+1} = X_t + V_t + \mu_t \tag{1}$$

with motion vectors for moving object V_t and prediction errors as μ_t. A prediction errors μ_t is calculated based on $|uI_x + vI_y + I_t|$ where I_x, I_y and I_t obtained from its partial derivative of x, y , and t, respectively. Thus, particle weights based on the likelihood function using the following observation model,

$$P(Z_t|X_t) = P(Z_t^{int}|X_t)P(Z_t^{mot}|X_t)^{O_t}P(Z_t^{trj}|X_t)^{1-O_t} \tag{2}$$

where Z_t consider intensity measurement Z_t^{int}, motion measurement Z_t^{mot} and trajectory measurement Z_t^{trj}, respectively.

Intensity measurement uses SSD to find out the similarity between the object initialization and the object position candidate at current time. A assumption state that the clutter is uniformly distributed over the area around x_t, the likelihood intensity is expressed in

$P(Z_t|X_t) = q_0U + C_N \sum_{j=1}^{J} q_j N(r_t, \sigma_t)$ with the number of candidates that match from the object initial based on SSD $q_0 + \sum_{j=1}^{J}$ where C_N is a normalization factor and q_j as prior probability. In its application, Huang uses $q_j = 0,05$.

Motion likelihood considers change in object position in the frame-t with the frame-$(t-1)$ as (x_t, y_t) to the average of object position change in the previous frames $((\Delta\bar{x}_t, \Delta\bar{y}_t)$. This is expressed as

$$\Delta\bar{x} = \sum_{S=t-k}^{t-1} \frac{x_s - x_{s-1}}{k}, \Delta\bar{y} = \sum_{S=t-k}^{t-1} \frac{y_s - y_{s-1}}{k} \tag{3}$$

So, the object position change from current frame to the object change on average is written as $d_{mot}^2 = (|\Delta x_t| - \Delta\bar{x}) + (|\Delta y_t| - \Delta\bar{y})$ with $t > 1$. Based on that, Motion likelihood is expressed as

$$P(Z_t^{mot}|X_t) = \frac{1}{\sqrt{2\pi}\sigma_{mot}} e^{\frac{d_{mot}^2/F}{2\sigma_{mot}^2}} \tag{4}$$

This likelihood becomes a consideration if object motion can be detected by motion estimation.

Calculation of trajectory likelihood as

$$P(Z_t^{trj}|X_t) = \frac{1}{\sqrt{2\pi}\sigma_{trj}} e^{\frac{d_{trj}^2/F}{2\sigma_{trj}^2}} \tag{5}$$

considers the closest particle distance to the estimated trajectory. For example, an estimated trajectory is expressed in a polynomial $y = \sum_{i=0}^{m} a_i x^i$ with a_i as the polynomial coefficient, then the closest distance between the particles with the estimated trajectory is calculated using $d_{trj} = |y - \sum_{i=0}^{m} a_i x_i|$. Whereas, $F = \lambda_f^{t_0}$ is the object effect ratio where the motion is not detected to the estimated trajectory. And, t_0 is the number of frames whose object motion cannot be detected by motion estimation. If more objects can't be detected, the reliability of the estimated trajectory tends to decrease. And, the value of λ_f follows Huang's suggestion, $(0 < \lambda_f < 1)$.

One of the ways if the object motion isn't detected $O_t = 0$ is a constructing a estimated trajectory. Calculation of the estimated trajectory utilize a prediction projection. A prediction is built based on an object motion in the previous frame expressed as

$$\hat{X}_{cur} = X_i + (X_i - X_j)\frac{cur - i}{i - j}, i > j \tag{6}$$

with $(i > j)$. Then, the results will be projected using

$$X_{cur} = (1 - \lambda_f^{t_0})\hat{X}_{cur} + \widetilde{X}_{cur}\lambda_f^{t_0} \tag{7}$$

by considering $\lambda_f^{t_0}$

2.2. Discriminative correlation filter

DCF for object tracking studies an optimal correlation filter between object initialization on the current frame to find out where object position is located. At the beginning of its development, DCF is only limited to estimate translational motion. However, recently DCF uses multidimensional feature representations included in object tracking as proposed by [4][9][10]. In his paper, [4] reviews the utilization of 1-dimensional filters as scale estimation, 2-dimensional filters as a translational estimation, and 3-dimensional scale filters form for combine scale estimation and translational estimation. The results of that approaches isnt only reliable for

translational motion estimation but also accurate in the object shape within scale variation change.

The learning process of multidimensional correlation filters usually consists a several patch of object appearances centered around the object. Then, a number of image patch in the form of $f_1, f_2, ..., f_d \in R^d$ will be a training sample with the desired correlation output as $g_1, g_2, ..., g_d$ with a filter h_t^l. In the correlation filter, it becomes the basic material to learn the correlation filter h which consists of one filter per channel feature in the frequency domain or Discrete Fourier Transform (DFT). This is obtained by minimizing the sum of squared errors of the h_t filter at current time. Functions f_1, h_1 and g_1 are in the same dimension and at the same time. The addition of a small value with $\lambda > 0$ in the SSE aims to improve the stability of the filter because the frequency is very low or zero which caused by division to zero [7].

$$\epsilon = \left\| g - \sum_{l=1}^{d} h^l * f^l \right\| + \lambda \sum_{l=1}^{d} \left\| h^l \right\|^2 \tag{8}$$

The star $*$ shows the element-wise multiplication using the convolution theory. The desired correlation output g is usually arranged in the form of a Gaussian function with a standard deviation parameter [7].

An optimal filter can be obtained by minimizing output errors in all training. To obtain a corresponding filter with a small error, the SSE can be transformed to the frequency domain using Parseval's formula [11]. The solution for formalizing an optimal filter is written in [4] as

The capital letters above indicate the existence in the Discrete Fourier Transform (DFT) domain. And, the capital bar shows the complex conjugation.

A robust filter in practice is obtained from learning on the number of image patches in each different time change $\{f_j\}_t^l$ because of each change in time, the object will also suffer changes in conditions and characteristics. Galoogahi [9] successfully solve this problem by modeling it in a linear least square problem. However, the process requires high computation time to solve each element in a filter with a linear equation of size d d. However, in overcoming this problem, Daneljan compiled a robust approximation by updating the H_t^l correlation filter and its denominator which is employed separately as

$$A_t^l = (1 - \eta)A_{(t-1)}^l + \eta \bar{G}_t F_t^l, l = 1, ..., d \tag{9}$$

$$B_t^l = (1 - \eta)B_{(t-1)}^l + \eta(F_t^k F_t^k) \tag{10}$$

with η as learning ratio parameters.

The object position in the new frame is determined based on the results of z_t extract from a considered transformation area. Then, the correlation score in the DFT domain is calculated as

$$Y_t = \frac{\sum_{l=1}^{d} \bar{A}_{t-1}^l Z_t^l}{B_{t-1} + \lambda} \tag{11}$$

z_t corresponds with image patch centered around the area where the predicted object position is. The object location is obtained from the maximum value of the correlation score with the inverse Y_t. So, the result is the coordinates where the center of the estimated object is located. as mentioned in [4] [Daneljan2], learning scale and translation filter are estimated separately. Learning correlation filters to estimate scale changes is developed into a 3-dimensional filter. The $f_{(t,scale)}$ training sample based on a feature extract using patch size variables centered around the object. For example, PR, is a representation of the size object in the current frame .An image I_n centered around the object will be extracted with the size $a^n P \times a^n R$ for each $\in [-\frac{(s-1)}{2}], ..., [\frac{(s-1)}{2}]$ as the scale filter size and a as the scale feature of the layer feature.

The extracted results from the training sample as $f_{(t,scale)}(n)$ are collected to the d-dimension descriptor feature of I_n. Then, the image representation using HOG by [12] aims to learn filters based on the d-dimension descriptor feature that is obtained from the training sample. In the [4] [10] study, filter learning using HOG can significantly improve object tracking accuracy.

3. Tracking approach

In this section, we will describe our combination based on discriminative correlation filters and adaptive particle filters in accordance with the above description. In its application, we use two tracking approach where both approaches have their reliability and weakness. Based on an understanding of the characteristics of the two tracking approaches, the weakness of a tracking approach will be overlap with the other tracking approaches and vice versa. So, the utilization of both will make tracker more reliable to overcome various tracking problems. However, this will make a double-edged knife where the utilization of two approaches together will make enlarged computing time. Besides aiming at making high accuracy, we also expect tracking which has a low computation time so that it can be implemented in real-time. To fulfill this expectation, we also devise an alternative solution in the leap-tracking form to reduce computational time.

3.1. Combination approach based particle and correlation filter

Here, we will discuss how an object tracking approach combined together between the adaptive particle filter (APF) and the discriminative correlation filter (DCF). We begin from the initialization at $t = 1$ where its the first of all step to obtain all information and parameters related to the tracked object. Furthermore, we can also apply evenly histogram distribution to improve image quality without losing native information. These efforts will make object more contrast than the surrounding objects. Then, after this step and so on we will find a tracking which begun from motion estimation to obtain a motion vector [13].

The main framework of our combination refers to the APF as the main tracker and compose DCF to support in the main framework. DCF is very reliable to estimate translational and changes in object shape especially in scale change. So, DCF will be embedded in the state transition and tasked to estimate object motion from translational and scale change. An estimate of object motion is obtained from the difference in the vector direction between the position of the object in a frame - $(t-1)$ and the spectrum peak position of Y_t in the spatial domain. Meanwhile, the APF as the main framework in tracking refines the motion object caused by an abrupt object change or the fast object motion and predict object position if the object's motion isnt detected or occluded. In its utilization, both will run in series where DCF will work after the first frame $(t = 1)$.

There are two states of motion estimation results i.e object motion detected or not detected. If object motion is detected, we utilize particle updates and errors as $X_t \sim N(X_{(t-1)} + V_t, t)$ based on the intensity measurements and motion measurements, respectively. Estimated errors in this condition are obtained from partial derivatives as described in section 2.1 . Each particle from this process is selected in the resample step which aims to removing particles with low weights. The results of resample are used to determine the object position with a weighted average of all particles $N(\widetilde{X}_t^{(i)}, 1/N)$. Otherwise, if the object motion is not detected by object motion estimation, then in the resample step, the particle with the highest weight will be selected as a candidate for the object position. The particle will become a reference to construct an interpolation as a projection of the predicted position of the object $(\hat{X}_{cur})$. The result in the form of X_{cur} is an updated position where the object position in the current frame is located. Then, if current frame is the end of an image sequence then tracking is complete. Otherwise, the process continues to the motion estimation step where DCF will be assigned.

There are three main processes in DCF by [4], namely translation estimation, scale estimation and filter update. In the DCF scenario, translation estimation work first before scale estimation
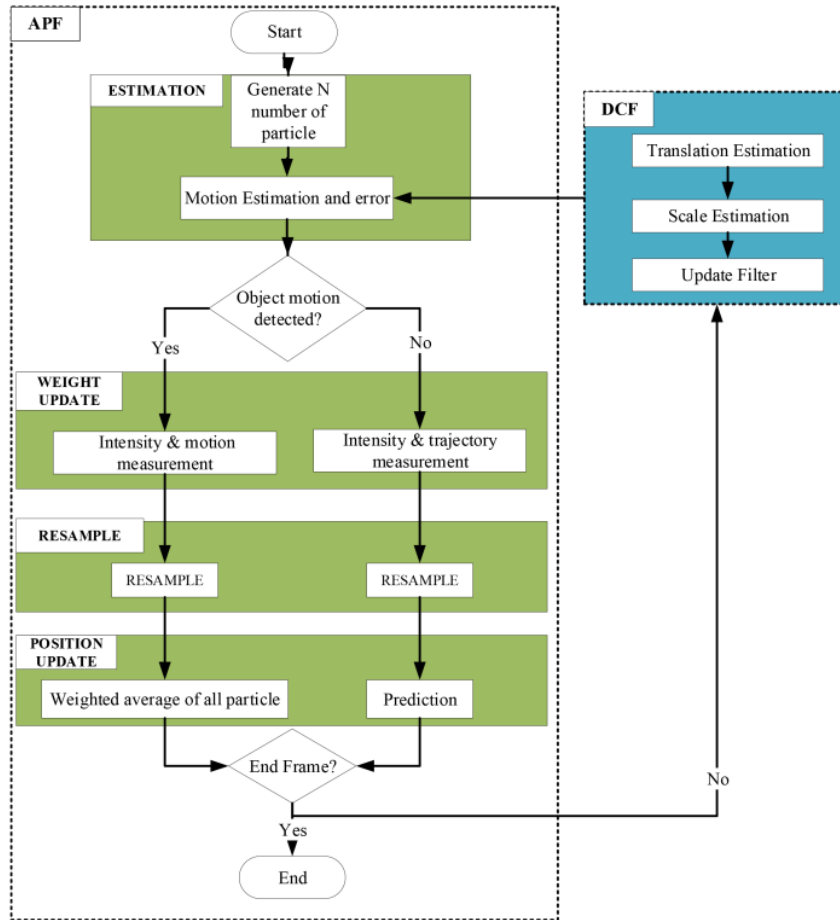
Figure 1. Blok Diagram of the combination between APF and DCF

to obtain the updated object position in the current frame. After that, scale estimation adjusts to the position where the object is located by maximizing the correlation score (y_t). Also, the learning filter between translation and scale work separately to construct each feature representation based on the appropriate search area. This will reduce computational time, especially in learning scale filters [4]. Then, update filter use (10) which is built from the filter correlation in the current and previous frames taking into consideration the learning ratio (η).

3.2. Leap-tracking

In general, object tracking process through frame by frame sequentially as it run in each frame of an image sequence. It shows that every frame of the image sequence always goes through the tracking process to search where the object is located. This process is a part which requires a lot of computation time, especially if the object search area is wide. However, if tracking works
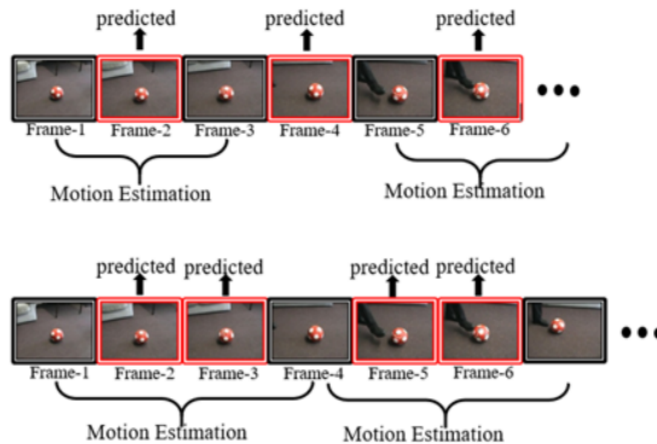
Figure 2. Visualization of the leap-tracking

by leaping (1-frame, 2 frames, or more) it will be enough to reduce the computation time of each frame that is leaped.

The number of leaped frames depends on the object condition. The important to be considered in this condition is objects do not run quickly and a sufficient framerate of the video from produced frame in every second (fps). Fast object motion or low FPS usually requires a wide search area so that computing time swell or failure in tracking accuracy. However, the large search area for fast object motion doesn't always make a large computation time. And, it all depends on the reliability and uniqueness of the tracking approach in overcoming fast object motion based on computational time. Meanwhile, in the leaped frame, the object position will be predicted in this case to adjust the prediction model of the adaptive particle filter which has reviewed above.

4. Result and discussion

We also follow a quantitative evaluation procedure using distance precision (DP), center location error (CLE) and overlap precision (OP) on image sequence from [14]. The first evaluation, CLE is calculated according to the average distance of Euclid between the estimated object position and the true object position. While on the second evaluation, the DP is calculated according to the relative number of frame in tracking where the deviation of the estimated object position is limited or not more than a certain threshold. We apply the same threshold value as [6] [14] which is 20 pixels in means to test our combination. The third evaluation, OP is calculated based on the percentage of the bounding box which exceeds the threshold of 0.5 in [4] and corresponds to the PASCAL evaluation criteria. Generally, the evaluation is useful to measure how reliable the tracking accuracy for the scale estimation of the object contained in the bounding box according to the true object. Moreover, a measurement in fps is used to measure the tracking speed. Finally, we also build a video by ourselves to test how it applies in real conditions.

To apply our combination, we use the same parameters as [8] for APF and [4] for DCF. The number of particle parameters in APF affects the tracking performance. Usually, the greater number of particle parameters tends to make the tracking accuracy better. However,

Table 1. Comparison of our combination and DCF on 15 benchmark tracking datasets in overlap precision (OP)(in percent), distance precision (DP) (in percent) and center location error (CLE)(in pixel)

	Basketball			Boy			Bolt		
	CLE	DP	OP	CLE	DP	OP	CLE	DP	OP
APF+DCF	9.91	98.9	63.2	2.28	100	100	63.6	68.9	65.1
DCF	39.8	76	48.8	3.21	100	100	410	2.29	2.29

	Couple			Crossing			Deer		
	CLE	DP	OP	CLE	DP	OP	CLE	DP	OP
APF+DCF	14.7	69.3	56.4	2.08	100	100	5.85	100	100
DCF	129	10.7	10.7	1.64	100	100	16.7	78.9	78.9

	Freeman1			Football			Ironman		
	CLE	DP	OP	CLE	DP	OP	CLE	DP	OP
APF+DCF	7.26	96.9	81.6	15.6	80.1	63.3	189	13.9	13.3
DCF	7.05	95.4	81.9	3.21	79	62.2	410	2.29	2.29

	Jogging			Matrix			Shaking		
	CLE	DP	OP	CLE	DP	OP	CLE	DP	OP
APF+DCF	14.1	73	63.8	98.3	17	17	7.09	100	100
DCF	149	18.9	18.2	73.6	19	13	94.7	1.37	1.37

	Subway			Trellis			Walking		
	CLE	DP	OP	CLE	DP	OP	CLE	DP	OP
APF+DCF	4.27	100	96.6	3.45	100	97.7	2.73	100	99.8
DCF	147	24,6	22.3	3.45	100	97.7	1.65	100	99.8

as a consequence, the tracker has a greater workload and computational time. Therefore, the number of particle parameters adjusts to problem level and tracking efficiency. In a different way, DCF by [4] is set to use fixed parameters in each processing step.

4.1. Quantitative evaluation

Table 1 shows the results of our testing discussed in section 3 to 15 image sequence from benchmark tracking datasets with various tracking problems. These table presents test results from tracking accuracy with OP, DP CLE, and tracker speed in fps. Based on this presentation, our combination can overcome the tracking problem as success DCF by [4] such as in boy, crossing, trellis, walking, and walking2. The same thing is also found in Ironman, and Matrix where there is no significant difference in accuracy between both. However, other case occur from other image sequences where there is an significant increase in accuracy with our combinaation.

Based on the test result in Table 1, Figure 3 shows some diagram built from the tracking result performance of our combination. This diagram is built on the level of tracking accuracy in our combination on 15 benchmark tracking datasets grouped and marked with different colours as Table 1 according to how much improved performance was achieved and success for implement leap-tracking. The performance of our combination is obtained from the average difference between DCF and our combination for each image sequence with a similar color group. As

Table 2. The comparison of computation time on 15 benchmark tracking datasets in frame per second (fps)

	Basketball	Boy	Bolt	Couple	Crossing	Deer	Freeman1
APF+DCF	3.41	13.6	19.7	7.49	39.1	6.4	45.1
DCF	1.77	34.3	21.1	22.7	38.4	7.22	51.9

	Football	Ironman	Jogging	Matrix	Shaking	Subway	Trellis	Walking
APF+DCF	42.7	14	19	17	13	30.3	10.8	30.4
DCF	22.7	16.9	8.77	23.8	10.9	27.9	6.78	28.3

(a)

(b)

(c)

(d)

Figure 3. The performance of our combination

in Figure 3, There is a significant increase in the group of image sequence with red and black respectively for 36.002 percent and 27.416 percent in Distance Precision (DP) , 42.302 percent and 20.617 percent in Overlap Precision (OP). In the same group from image sequence, Center Location Error (CLE) in red and black with our combination can suppressed respectively 25 pixels and 50.31 pixels. But, when referring to the success rate in computation time efficiency, the group of the video in black can be a reference for how much improved performance of our combination and show an increase in accuracy for 29.416 percent in Distance Precision (DP), 20.617 percent in Overlap Precision (OP), and 50.31 pixels in Center Location Error (CLE).

In addition, Table 2 also shows the tracking speed in fps. As a statement in the above where there is a consequence if the tracking approach is assigned together i.e higher computing time. This case is found in Basketball, Bolt, Couple and Deer where every second our combination is only able to process 4 - 8 frames with take into account high tracking accuracy. However, our combination can pursue even exceed DCF speed in other image sequence. Figure 4 shows the
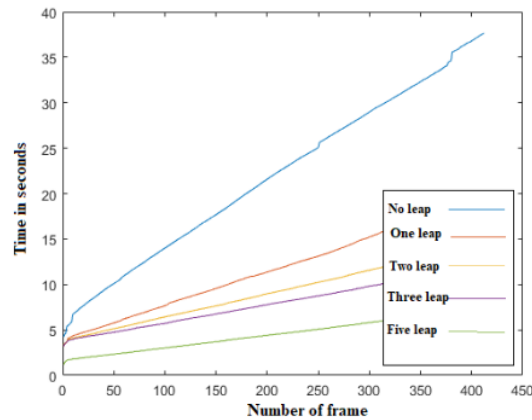
Figure 4. Comparison of computation time from a number of different frame leaps

comparison of computation time from a number of different frame leaps on walking. The next section review qualitatively tracking performance evaluations how this situation can occur .

4.2. Qualitative evaluation

This section reviews the qualitative evaluations based on the test results in Table 1. There is a similarity between the test results from DCF with our combination. The difference between both of them occurs on the computational time in our combination which tends to have higher computing time as happened to Boy. Usually, The problems can be overcome with utilize leap-tracking to reduce computing time. However, leap-tracking isnt suitable in conditions such as Table 1 in blue colors because of fast motion (FM), motion blur (MB) and Illumination problem. If our combination is assigned to Ironman and matrix with the same problem, there is no significant increase in accuracy compared to ordinary DCF. The same thing happens to boy where leap-tracking is constrained by abrupt object motion and blur (MB).

In other cases, although it cannot use leap-tracking, there are several image sequences which tracked with better accuracy. This is shown in Table 1, which is red, including deer, couple, bolt and basketball. DCF is very good to overcome some tracking problem but it's constrained in fast motion (FM), motion blur (MB), deformative (DEF) and background clutter (BC). DCF is constrained by BC and MB in deer so that in the frame 26, tracker drift from the true object position. But, in the frame 37, the object is tracked again accurately. DCF is also constrained by FM in couples from the frame 33 into the last frame. The DCF problem in bolt, couple, and basketball occur in the deformation because of the shape and direction change of the object. The problem can be corrected by APF work result so that deer and basketball get very good accuracy even though basketball obtain 63.2 percent in OP due to out of plane rotation (OPR) problem. Meanwhile, Bolt and Couple also get more accuracy after correction APF even though it only obtain 68.9 percent and 69.3 percent in DP respectively with the same problems.

Table 1 in black colors, is a tracking result that can be applied to leap-tracking so it can reduce computational time. Even though its successfully leap tracking, tracking results can still show good performance such as crossing, trellis, and walking. Meanwhile, among the 15 testing in the image sequence, shaking show the best results where our combination have significant increase in accuracy from 1.37 percent in DP and OP into 100 percent in both. In addition,
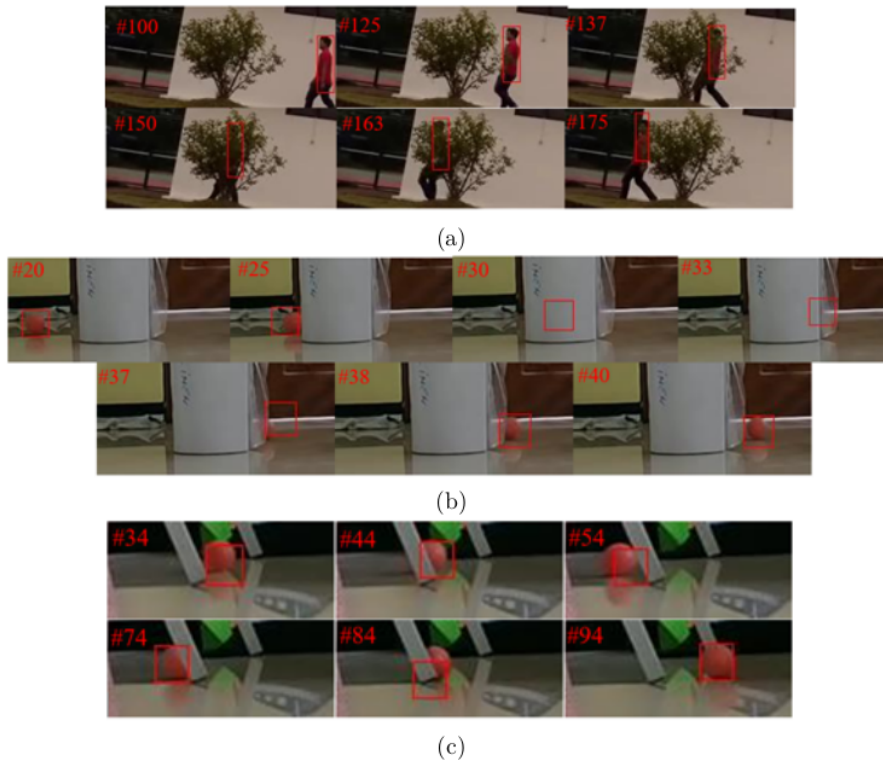
(a)

(b)

(c)

Figure 5. The Result of 3 image sequences by ourselves

our combination also successfully suppresses CLE from 94,7 pixels into 7,09 pixels and reduces computation time thereby increase processing frame per second from 10,9 fps into 13 fps. From that images sequence, the tracking failure of the DCF is located at the background clutter (BC) problem in the beginning frame so in the next frame into the end, tracking suffers accuracy drift.

Finally, Table 1 in black, also contains the tracking result with deformation (DEF), background clutter (BC), occlusion (OCC), where our combination significantly successfully overcome that problems. This problem, especially in full OCC, is a weakness of correlation filter-based tracking [15] which continuously identifies within the search area where the corresponding object is located. Sometimes in some cases, a small part which covers an object is identified as the tracked object so that if there is a full occlusion, tracking will suffer an error in object identification[16]. This case occurs when the visibility of the tracked object becomes limited or completely miss while the object has not left the camera's viewpoint [16]. From this problem, leap-tracking is an alternative solution to avoid misidentification by substituting into a prediction in APF.

Figure 5 shows the results of an image sequence by ourselves with focusing on a full occlusion problem. Figure 5(a) shows repeatedly occlusion and full occlusions in Figure 5(b). Meanwhile, Figure 5(c) also shows a case of occlusion with zig-zag moving objects. The number of leaped frames depends on how many objects are blocked in a certain number of frames. As in jogging, our combination requires a leap with 10 frames to avoid occlusion but in the end of images sequence, tracking drift due to OPR.

5. Conclusion

This article presents a combination of tracking using APF and DCF. We implemented leap-tracking to reduce the high computation time through this combination in fps. Leap-tracking also becomes an alternative solution to overcome full occlusion and avoid object identification errors. Our approach employs both in series where DCF works as an estimation of object motion. APF is assigned to improve the estimation results from the object motion with a certain number of particles as a representation of the candidate's center point.

The testing of our combination use 15 benchmarks with various tracking problems. The test results show that our combination have improved performance significantly for 29.416 percent in Distance Precision (DP), 20.617 percent in Overlap Precision (OP), and 50.31 pixels in Center Location Error (CLE). In addition, the reducing of tracking procedure efficiently increase tracking speed without reducing accuracy. An effort to avoid misidentification in the occlusion problem shown in 3 image sequences by ourselves.

References

[1]   Yilmaz Alper, Javed O and Shah M 2006 Object tracking: A survey *Acm computing surveys (CSUR)* **vol 38** 4 13.
[2]   Swalaganata G, Sulistyaningrum D R and Setiyono B 2017 Superresolution imaging applied to moving object tracking *Journal of Physics: Conference Series* **893** 1 012062.
[3]   Habibi Y, Sulistyaningrum D R, and Setiyono B 2017 *AIP Conf. Proc* vol 1867 1 p 020024
[4]   Danelljan Martin Hager G Khan F and Felsberg M 2014 *British Machine Vision Conf* ed Valstar M, French A and Pridmore T (Nottingham : BMVA).
[5]   Danelljan M Shahbaz F K Felsberg M and Van D W J 2014 *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* pp 1090-97.
[6]   Henriques J, Caseiro R, Martins P and Batista J 2012 *European conference on computer vision* (Berlin : Springer) pp 702-15 .
[7]   Bolme D S Beveridge J R Draper B A and Lui Y M 2010 *IEEE Computer Society Conf. Computer Vision and Pattern Recognition* (San Francisco:IEEE) pp 2544-50
[8]   Huang Y and Llach J 2008 *Int. Conf. on Audio Language and Image Processing* pp 357-62
[9]   Galoogahi HK Sim T and Lucey S 2013 *Proc of the IEEE Int.Conf. Computer Vision* p 3072-79
[10] Henriques J, Caseiro R, Martins P and Batista J 2014 *IEEE transactions on pattern analysis and machine intelligence* vol 37 3 pp 583-96.
[11] Danelljan Martin Hager G Khan F and Felsberg M 2016 Discriminative scale space tracking *IEEE transactions on pattern analysis and machine intelligence* **39** 8 1561-75
[12] Felzenszwalb P F, Girshick R B, McAllester D A and Ramanan D 2009 Object detection with discriminatively trained part-based models *IEEE Transactions on Pattern Analysis and Machine Intelligence* **vol 32** 9 162745.
[13] Sulistyaningrum D R, Setiyono B, and Rizky M S 2018 Tracking and people counting using particle filter method *Journal of Physics Conf. Series* **974** 1 012059.
[14] Wu Y. Lim J and Yang M H 2013 *Proc. IEEE Conf. Computer Vision and Pattern Recognition* pp 2411-18
[15] Zhao 2017 *AIP Conf. Proc* vol 1839 1 p 020168.
[16] Lee B Y, Liew L H, Cheah W S and Wang Y C 2014 *IOP conf. series: earth and environmental science* **18** 1 012020.

PRIMARY SOURCES

1	Submitted to School of Business and Management ITB Student Paper	**4**%
2	backend.orbit.dtu.dk Internet Source	**1**%
3	liu.diva-portal.org Internet Source	**1**%
4	Yu Huang, J. Llach. "Tracking the small object through clutter with adaptive particle filter", 2008 International Conference on Audio, Language and Image Processing, 2008 Publication	**1**%
5	export.arxiv.org Internet Source	**1**%
6	Makoto Mizukami, Takashi Yoshizawa Takashi Yoshizawa, Isamu Sato Isamu Sato. "High-Precision Head Positioning Using Multiple Offset Beams", Japanese Journal of Applied Physics, 1993 Publication	**<1**%

7 web.archive.org
Internet Source

<1 %

8 Qiang Fan, Gangbo Sun, Erbo Zou. "Tracking of small infrared targets based on convolutional neural network", Journal of Physics: Conference Series, 2020
Publication

<1 %

9 Alper Yilmaz. "Object tracking", ACM Computing Surveys, 12/25/2006
Publication

<1 %

10 Lecture Notes in Computer Science, 2015.
Publication

<1 %

11 Yan-Chun Chen, Wei-Chan Weng, Shi-Wei Lin. "A High Reliability 3D Object Tracking Method for Robot Teaching Application", IOP Conference Series: Materials Science and Engineering, 2019
Publication

<1 %

12 dokumen.pub
Internet Source

<1 %

13 Submitted to SUNY Brockport
Student Paper

<1 %

14 Xincong Yang, Heng Li, Ting Huang, Ximei Zhai, Fenglai Wang, Chen Wang. "Computer-Aided Optimization of Surveillance Cameras Placement on Construction Sites", Computer-

<1 %

Aided Civil and Infrastructure Engineering, 2018
Publication

15 Galandaru Swalaganata, Muniri, Yessi Affriyenni. "Moving object tracking using hybrid method", 2018 International Conference on Information and Communications Technology (ICOIACT), 2018
Publication

<1 %

16 Huanlong Zhang, Zeng Gao, Xiaoyang Ma, Jie Zhang, Jianwei Zhang. "Hybridizing Teaching-Learning-Based Optimization With Adaptive Grasshopper Optimization Algorithm for Abrupt Motion Tracking", IEEE Access, 2019
Publication

<1 %

17 jivp-eurasipjournals.springeropen.com
Internet Source

<1 %

18 Fujian Chen, Weixin Xie, Ting Xia. "Target Tracking Algorithm Based on Kernel Correlation Filter with Anti-occlusion Mechanisms", 2020 15th IEEE International Conference on Signal Processing (ICSP), 2020
Publication

<1 %

19 Lidan Li, Yi Zhang, Qianya Guo, Qing Lin, Jing Zhou, Yanyong Wang, Bo Zhao. "Real-time Object Tracking Method based on Multi-Core DSP", Journal of Physics: Conference Series, 2021

Publication

20   koreascience.or.kr
      Internet Source   <1%

21   dokumen.dinus.ac.id
      Internet Source   <1%

22   B Y Lee, L H Liew, W S Cheah, Y C Wang.   <1%
      "Occlusion handling in videos object tracking:
      A survey", IOP Conference Series: Earth and
      Environmental Science, 2014
      Publication

23   vdoc.pub
      Internet Source   <1%

24   www.biorxiv.org
      Internet Source   <1%

25   www.mdpi.com
      Internet Source   <1%

26   "Computer Vision – ECCV 2016 Workshops",   <1%
      Springer Science and Business Media LLC,
      2016
      Publication

27   Jian Wei, Feng Liu. "Online Learning of   <1%
      Discriminative Correlation Filter Bank for
      Visual Tracking", Information, 2018
      Publication

28   Lei Song. "A Comprehensive Study on Backup-   <1%
      Bandwidth Reprovisioning After Network-

State Updates in Survivable Telecom Mesh Networks", IEEE/ACM Transactions on Networking, 2008

Publication

---

29 Rawan Fayez, Mohamed Taha Abd Elfattah Taha, Mahmoud Gadallah. "Occluded Object Tracking System (OOTS)", International Journal of Service Science, Management, Engineering, and Technology, 2020

Publication

<1%

---

30 umpir.ump.edu.my
Internet Source

<1%

Exclude quotes	Off		Exclude matches	Off
Exclude bibliography	On			