International Journal of Current Advanced Research
ISSN: O: 2319-6475, ISSN: P: 2319-6505, Impact Factor: 6.614
Available Online at www.journalijcar.org
Volume 7; Issue 5(G); May 2018; Page No. 12722-12728
DOI: http://dx.doi.org/10.24327/ijcar.2018.12728.2246

Research Article

TREE-BASED APPEARANCE MODEL FOR ROBUST VISUAL TRACKING

Bhuvaneswari R[1] and Ravi Subban[2]

[1]Mother Teresa Women's University, Kodaikanal, Tamilnadu, India
[2]Department of Computer Science, School of Engineering and Technology, Pondicherry University, Pondicherry, India

A R T I C L E I N F O

Article History:

Received 20th February, 2018
Received in revised form 20th March, 2018 Accepted 8th April, 2018
Published online 28th May, 2018

Key words:

Visual tracking, Bayesian model, Shallow Learning, Deep Learning

A B S T R A C T

In recent years, visual tracking become more popular in the field of computer vision. For effective visual tracking, the tracking method should be capable of separating the target object from the background accurately. While designing the model of visual tracking, several issues need to be considered. Some of the issues are occlusion, scale variation, rotation, motion blur, deformation and background clutter. In order to achieve effective visual tracking, a Tree-based appearance model is proposed. The proposed method characterizes the target in two levels: local level and global level. In local level, a group of local patches are utilized to map the target for adjusting the variations in appearance. In the global level, the target is mapped by double bounding boxes based on foreground and background. The interior box includes the target region and the exterior box includes both the target and background region around the target. The target object is encoded by two HSV color histograms and the drifts are suppressed during the tracking process. The position of the object is calculated by increasing the likelihood of the target using Bayesian method. The performance of the proposed method is evaluated and compared with shallow and deep collaborative model. The simulation results shows that the proposed method produces better results than shallow and deep collaborative model in terms of efficiency, accuracy and robustness.

INTRODUCTION

Visual tracking plays a vital role in computer vision [1]. Many researchers are carried out in visual tracking because of several practical applications. It is widely used from military applications to civilian applications like border surveillance, object tracking, behavior analysis, Human Computer Interaction (HCI), etc. [2]. The difficulty of the tracking system based on various factors like amount of prior information known about the target object and number of parameters monitored (location, scale). Existing tracking methods works well and produce effective results only in simpler situations, i.e. slow motion, less occlusion, etc [3]. Tracking generic objects remains a very challenging task because of the presence of blur, rotation, fast motion, occlusion, scale variation, pose change and background noise. These challenges are illustrated in Fig. 1. So, a new effective tracking model is needed to solve these issues. A traditional tracking model comprises of three units: appearance model, motion model and search strategy [4]. The appearance model calculates the probability of the object of interest at some particular location. The motion model interconnects the location of the object with respect to time.

Corresponding author:* **Bhuvaneswari R
Mother Teresa Women's University, Kodaikanal, Tamilnadu, India

The search strategy model finds the more appropriate location of the object in the current frame. An appearance modeling is used to develop a mathematical model for the identification of objects, especially in visual tracking [5]. Generally, there are two types of trackers which include generative tracker and discriminative trackers. Generative trackers find the target which looks same as target object. Thismethod uses templates, subspace or inference methods. Discriminative tracking method uses binary classification problem to differentiate the target from the background.

The tracking methods can also be classified into two methods with respect to representation mode. The two methods are holistic method and parts-based method. The holistic method captures larger objects by mapping the appearance of the target with global cues [6-8]. For local visual cues, holistic method fails to produce better results. On the other side, parts -based method mapped the target appearance with local patches, which encodes the spatial information. These patches are not connected tightly to some extent of spatial deformation. This method can be used for short term tracking and produces better results than holistic method. This method works well in presence of partial occlusion. Due to non-consideration of the entire information of the object, it fails to track objects in presence of background clutter and motion blur. It is clear that both holistic and parts based method considers only a part of visual cues, either global or local cues. These methods use the information within the object and ignoring the textual cues. The existing tracking methods fail to produce effective results in complicated situations.

Higher value of reliability represents higher possibility of a local patch corresponding to the object, and vice versa.

Fig 1 Challenges during tracking in real-world environments, occlusion (woman), abrupt motion (shaking), illumination change (carDark), pose variation(Bird) and complex background (board).We use blue, green, black, yellow, magenta, cyan and red rectangles to indicate the tracking results of the IVT [9], ASLA [10], OSPT [11],CT [12], Struck [13], DLT [14] and the proposed method, respectively.

In order to visually track the object in complicated situations, a new tracking approach calledTree-based appearance method is proposed. Tree-based appearance model uses local and global levels under the Bayesian framework. The working flow of the proposed method is shown in Fig 2. In Fig 2, various image regions are indicated by several colors bounded by box. In the local level model, the desired image region (marked by yellow color) is partitioned into smaller grid shaped local image patches, indicated by green color. A bigger region (indicated by blue color) is used to determine the reliability. The bigger region also has the same origin and is based on the size of the target. Additionally, the bigger region is partitioned into equal sizes as patches in the target region, and every patch is indicated by a green rectangle. In the local level, local patches are used to adjust the variations in appearance. To determine how closely a patch pertaining to the target, likelihood is defined as the weighted sum of reliability index and stability index. Reliability index is defined as the closeness of a local patch respective to the target object.

Stability index defines the variation degree of a local patch in successive frames.

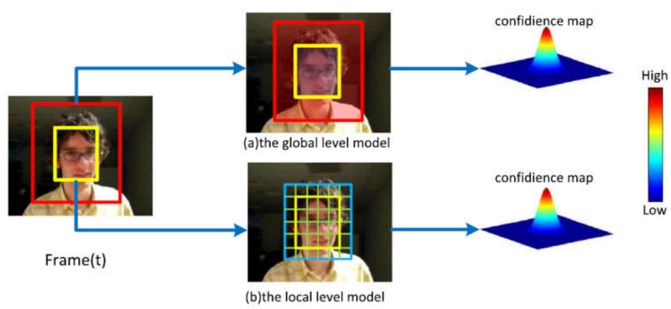

Fig 2 The flow of the proposed Tree-based appearance model

Higher stability index represents lesser changes in the local patch and vice versa. This gives the reliability of correctly identifying the target object. In the global level, the target is mapped by double bounding boxes based on foreground and background. The interior box includes the target region and the exterior box includes both the target and background region around the target. The target object is encoded by two HSV color histograms and the drifts are suppressed during the tracking process. With the help of these color histograms, the possibility of the pixel corresponding to a target can be estimated. Confidence maps can be generated by estimating the pixels at the global level. Using the confidence map of local and global levels, the candidate with maximum posterior probability under the Bayesian tracking framework can be observed. In Bayesian framework, the candidate samples are sampled stochastically depending upon the previous states with respect to Gaussian distribution. Every individual sample consists of a collection of patches and every individual is mapped as Gaussian distribution based on the temporal changes. Under proposed method, the samples and local patches are defined and the sample with highest likelihood score is chosen as next target state. With the integration of contextual information and local patches of the target, the proposed model achieves accuracy and robustness.

The contribution of the paper is summarized as follows. A Tree-based appearance modelis proposed by combining local and holistic models together. It uses contextual information of the target and spatial-temporal cues from the local patches. Next, likelihood criteria with reliability and stability indices are used to determine the possibility of the patch related to the target. This proves that reliability and stability indices are independent of partial occlusion, motion blur, etc.

The remainder of the paper is organized as follows. Related work is given in Section 2. Bayesian tracking framework is described in Section 3. The proposed tracking procedure is explained in Section 4. The proposed method is simulated and results are discussed in Section 5 and the paper is concluded in Section 6.

Related work

In the last decade, huge number of model is proposed in the field of target tracking. For the sake of simplicity, the existing models which are more relevant to the proposed method are discussed. Feature extraction is the important task in tracking process. In terms of representation mode, tracking models can be classified into holistic approach and local-based approach. Holistic approach considers that the target remains unchanged in successive frames. Some researches show that the holistic method achieves better results due to its simpler representation. An appearance model is proposed to adaptively select the color features [15]. This method can easily differentiate the target from its background. An online Hough-based method is presented to track non-rigid objects [16]. For tracking non-rigid objects, [17] proposes a method to identify the target object by reducing the Bhattacharyya distance from the color histograms of reference model and target bounding boxes. An online appearance modeling technique is developed by the utilization of sequential density approximation [18]. Another online method is proposed [19] to discriminate the target object by filtering the background by training a discriminative classifier. In the process of training, the tracking result is considered as positive sample and surrounded bounded boxes are considered as negative samples. But, these methods can be easily affected by the presence of noise. Based on correlation filters, a fast tracking method [20] is proposed in a latest benchmark (OTB) [21]. This method does not consider the factor of online model updating. This method is fails to produce poor tracking results in presence of occlusion for a longer time. In recent days, visual saliency is employed to attain accurate results because of its better performance in object detection [22]. Computational complexity is the major drawback of visual saliency. It is clear that holistic methods is capable of acquiring larger size images but fails to adapt the appearance variations in presence of occlusion and deformation.

Parts-based approach produces better results in presence of occlusion and deformation [23-25]. In some partially occluded situations, the rest of the visible part is capable of providing reliable cues for tracking. A star like appearance model [26] is proposed to improve the tracking accuracy and robustness using segmentation techniques. A new method is developed to present the structural information between the interior parts of the target. A graph-matching method is used to interconnect the adjacent frames. This method performs well in the presence of occlusion and deformation. A part based algorithm [27] is developed to investigate the interconnection between the whole object and local patches. They used two-stage training process to identify the target position. A discriminative ranking list based tracker is presented to restrain the distractor [28]. In this method, the target is represented by the patches of smaller and larger sizes. Smaller patches are employed to calculate the confidence of every individual input patch and larger patches are utilized to eliminate the untrustworthy smaller patches. [29] splits the bounding box into several patches and it chose only the significant patches to produce a precise foreground probability map. Most of the reviewed approaches are slower in producing results in benchmark (OTB). To eliminate this drawback, a novel tracking method is proposed which track objects based on multiple correlation filters [30]. It produces accurate tracking and robust in the presence of occlusion and deformation. Several approaches uses local patch to track the objects. For enhancing the tracking performance, background information is combined with the tracking models. A couple layer model [31] is proposed which represent the target by the combination of local and global appearance.

Though single tracking models are computationally less complex, it is not suitable for situations with occlusion and noise. Multi-object tracking methods handle occlusion effectively by using high level associations. But, handle occlusion effectively byusing high level associations. Multi-object tracking methods are computationally more complex than single tracking method. In order to visually track the object in complicated situations, a new tracking approach called Tree-based appearance method is proposed. Tree-based appearance model uses local and global levels under the Bayesian framework for efficient and robust tracking.

Bayesian tracking framework

The proposed method uses Bayesian framework for tracking purposes. X_i is the target state at time t, Y_i is the observation set from time 1 to t, $Y_{1:t}$ = { Y_1, Y_2,..., ,...,Y_t}. The posteriori probability is updated with Bayesian filter and formulated as:

$$p(X_t|Y_{1:t}) \approx p(Y_t|X_t) \int p(X_t|X_{t-1}) p(X_{t-1}|Y_{1:t-1}) dX_{t-1} \quad (1)$$

where $p(Y_t|X_t)$is the observation model which computes the similarity between the observation at the present estimated state and the givenmodel; and $p(X_t|X_{t-1})$is the transition model that determines the nextstate $X_t$based on the previous state X_{t-1}. It is important that the dynamic change between two successive states in the state space is usually modeled bythe Brownian motion.$p(X_t|X_{t-1})$) is defined as:

$$p(X_t|X_{t-1}) = N\ (X_t, X_{1:t-1}, \Psi) \qquad (2)$$

A diagonal covariance matrix Ψ is used, where the elements of the matrix are the variance of the state parameters X_i. With posteriori probability $p(X_t|Y_{1:t})$, the Maximum a Posteriori (MAP) estimate over the *n* number of samples at each time *t* is determined by

$$\widehat{X_t} = arg_{x_t^{(l)}} \max p\left(X_t^{(l)}|Y_{1:t}\right), for\ l = 1, \ldots, n \qquad (3)$$

In Eq. (3), $X_t^{(l)}$denotes the l^{th}sample of the object state X_t; $\widehat{X_t}$indicates the best sample of the object state, It defines the object configuration accurately for the given observations. Representx_t^i i = 1, 2, ...,*m* as the i^{th}local patch of the object state $\widehat{X_t}$at time *t* , and *m* is the total number of local patches. Then, the observation model can be equated as

$$p(Y_t|X_t) = p_{\mathfrak{Z}}(Y_t|X_t) \prod_{i=1}^{m} p_{\zeta}\ (y_t^i|x\ _t^i) \qquad (4)$$

Note that in Eq. (4), $p_{\mathfrak{Z}}(Y_t|X_t)$and $p_{\zeta}(y_t^i\ |\ x\ _t^i)$indicates the global andlocal likelihood, respectively; $y_t^i$is the i^{th}observation with respect tothe state of the i^{th}local patch $x\ _t^i$at time *t*. For simplicity, a patch's stateand observation are denoted as $x_t$andy_t respectively.

Tracking procedure

The Tree-based appearance modelis proposed for accurate and reliable tracking usingBayesian tracking framework. The working of the proposed model is shown in Fig 3. Initially, the tracker is propagated with the help of past results. Frame (t-1) is considered as the reference for the present time Frame (t) to estimate the target location. The solid yellow line represents the tracking output of highest likelihood. With the help of reference, a sequence of candidate image is created by Gaussian model. Every region of the image has the possible location of the object. The proposed system defines the object by Tree-based appearance model. Tree-based appearance model consists of two levels: local level and global level. In global level, the target is predicted using the contextual information. In local level, stability and reliability indices are used to predict the target state in local patches. Using these indices, sequence of confidence maps are generated in the local level. Likewise, sequence of confidence maps are generated in the global level. These confidence maps are employed to predict the location of the target object in the upcoming frame with higher posterior probability. The proposed method operates in three levels: location, model updating and occlusion handling.

Location

Tracking is mainly used to find the optimal state of the object in the series of images. At the starting frame, manual selection of target objects is done with the help of bounding box. The appearance model is obtained by partitioning the target and its neighboring region into local patches. A set of candidate rectangular regions are sampled byGaussian distribution. By equation (3), the new target state$(\hat{x}_t^c, \hat{y}_t^c, \hat{s}_t)$ can be calculated, where $(\hat{x}_t^c, \hat{y}_t^c)$ is the center coordinate and $\hat{s}_t$is the corresponding scale of the object. For searching the optimal state$\hat{X}_t$, two level search strategies is used based on the Gaussian distribution model to gradually approximate the high score region. The two levels of search strategy are coarse sampling and fine sampling. To adapt the various types of motions, the variance $\sigma_t$of the transition model is updated using (2). In the beginning, $l_l$candidates are sampled every time and best state$\hat{X}_t^1$ is achieved by increasing the likelihood score.

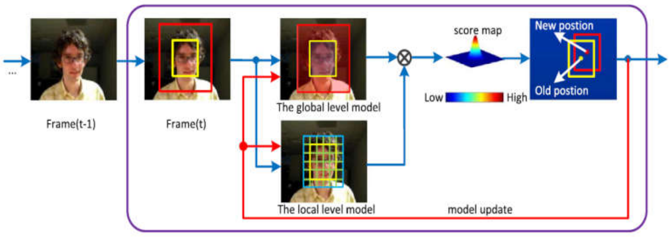

Fig 3 Flow of the proposed model

For larger search range, the variance is equated as

$$\sigma_{x,t}^1 = (W + H)/2 \qquad (5)$$

$$\sigma_{y,t}^1 = (W + H)/2 \qquad (6)$$

Next, l_2 candidates are re-sampled by $\hat{X}_t^1$and target is located by increased likelihood score. The variance is considered as a smaller value for proper tracking results, i.e. $\sigma_{x,t}^2 = \sigma_{x,t}^1/8$ and $\sigma_{y,t}^2 = \sigma_{y,t}^1/8$. As a result, this model works effectively for fast and slow motions.

Model updating

For robust target tracking, it is important to update the target model with respect to time. When new object location is identified, then the object model is updated. The local level model parameters$(\mu_l^i, \sigma_l^i)$are updated with the present patch descriptor$h_i(x_t)$and the default diagonal matrix $\sigma_0$is given as:

$$\mu_l^i = \lambda\mu_l^i + (1 - \lambda)h_i \qquad (7)$$

$$\sigma_l^i = \sqrt{(\lambda(\sigma_l^i)^2(1 - \lambda)(\sigma_0)^2 + \lambda(1 - \lambda)(\mu_l^i - h_i)^2} \qquad (8)$$

The global level is updated by linear interpolation which is given as

$$p_t(F|x) = \eta p(F|x) + (1 - \eta)p_{t-1}(F|x) \qquad (9)$$

Occlusion handling:

It is handled by using the similarity of patches between target and background [32]. The overview of the proposed system is Algorithm 1. At the global level, the target model is represented by the context information. At the local level, the target model is represented by the reliability and stability indices of local patches. The score map obtained from confidence scores of the two levels is used to infer the location of the object in the next frame.

Algorithm 1 Overview of proposed model

1. Input the previous tracking state$\hat{X}_{t-1}$
2. Sample $l_1$samples $X_t^{1,l} \sim N(X_t^1, X_{t-1}, \Psi)$, $\quad$ l: 1~ l_1
3. for each sample
4. Grid each sample to obtain $X_t^{1,l(i)}$, i: 1 ~ m
5. Compute $S_{rel}(i)$ and $S_{sta}(i)$
6. Compute the likelihood $p_{\zeta}(y_t|x_t)$ of every patch
7. Compute the probability $p_{cot}(F|x)$and$p_{dis}(F|x)$

8. Compute the likelihood $p_{\zeta}(Y_t|X_t)$
9. Compute the likelihood $p(y_t|x_t)$
10. end for
11. Determine the optimal state by max the score
12. Resample l_2 samples $X_t^{2,l} \sim N\left(X_t^{2}, \widehat{X}_t^{1}, \Psi\right)$, $l: 1\sim l_2$
13. Determine the optimal state $\widehat{X}_t^{1}$ by max the score $p(Y_t, X_t^{2})$
14. Update the local model by using Eq. (7) and Eq. (8)
15. Update the local model by using Eq. (9)
16. Output: New target state $\widehat{X}_t$

The equations to compute $S_{rel}(i)$, $S_{sta}(i)$, $p_{\zeta}(y_t|x_t)$, $p_{cot}(F|x)$ and $p_{dis}(F|x)$, $p_{\zeta}(Y_t|X_t)$ and $p(y_t|x_t)$ are referred from [33].

SIMULATION RESULTS AND DISCUSSION

To evaluate the performance of Tree-basedappearance model, an extensive experiment is conducted with 7 published challenging tracking sequences as well as visual benchmark to depict the advantages of Tree-based appearance model. The various sequences are Boy, Car4, Football, Freeman1, Shaking, Women and Walking. The proposed method is compared with 5 existing models includes IVT, CT, DFT, Shallow and deep collaborative model[33]. The sequences are filled with complex situations like motion blur, cluttered background, pose variations and partial occlusion. The experimental study is conducted in MATLAB. The performance metrics used to evaluate the proposed method are the center location error and the overlap rate. The centre equalization error (CLE) is defined as the Euclidean distance from the center of the tracking result (B_T) to the ground truth (B_G) of every individual frame. From B_G and B_T, overlap rate is calculated by

$$score = \frac{Area\ (B_T \cap B_G)}{Area\ (B_T \cup B_G)} \quad (10)$$

For precise tracking results, the overlap rate to be higher and average error to be lesser. At the time of initialization, some default parameter values are given. The simulation parameters are shown in Table 1.

Table 1 Simulation Parameters

Parameter	Value
Number of patches (m)	25, (n_w=5, n_h=5)
Number of samples, l_1	100
Number of samples, l_2	50
λ_r	1.0
λ_s	1.0
λ_e	2.0
η	0.1
λ	0.95
σ_o	0.03
N_e	4
N_s	4
thr	1.25

The results of the proposed method and existing methods are tabulated in Table 2 and Table 3 interms of center location error and overlap rate respectively. The better performance results are marked in red color.Higher value of overlap rate and lower value of center localization error represents accurate tracking. The comparison of the proposed method with existing models is shown in Fig 4 and Fig . Fig 4 shows the comparison of all models interms of average center location error for 7 benchmarks. Similarly, Fig 5 shows the comparison of all models interms of overlap rate for 7 benchmarks. From Table 2 and Table 3, the proposed method produces better results than existing model interms of center location error and the overlap rate for majority of sequences. The holistic approach IVT adapts to the appearance variations of the target by learning the appearance online.This is suitable for some situations in presence of noise like illumination variations, background clutter and distractor. But, it fails to provide better results in presence of occlusion. It does not have the ability to track non-rigid objects. DFT and the proposed method works well for Woman sequence.

For the Football and Boy sequences, the target supposed to fast motion in cluttered background. Most of the objects appear same as target in the Football sequence. Several methods tend to shift from desired target to other objects when they become closer to each other. The proposed method and SM accurately track the whole sequence. IVT, CT and DFT fail to track the objects with abrupt motion. DM tracks well but drifts at some time instances.For Boy sequence, proposed method tracks the target effectively in the whole sequence. SM and DM track the object effectively but it fails to track at some instances. IVT, CT and DFT leads to poor performance. From these observations, the proposed method handles abrupt motion effectively by multiple grid-like patches. Every individual patch represents stationary part of the object. The combination of local patches covers the entire structure of the target. Hence, the proposed method holds the interior details of the target in the tracking process.

The Freeman1 sequence has variations in scale and pose. Tree based method tracks the object in the entire sequences. SM and DM tracks effectively but varies from the target at some instances. IVT, LT and DFT fail to track the target accurately. These results shows that the proposed method is suitable for sequences with can handle pose variation. This is achieved by the utilization of multiple patches and the context information concurrently.

The walking sequence has scale variations. The proposed method achieves effective tracking. IVT also achieved closer results as Tree based method. Other existing methods LT, DFT. SM and DM lose its ability to track target to scale variation.

Car4 sequence contains significant variations in scale and illumination. IVT, SM, DM and proposed method tracks the car in the entire sequence. Shaking sequence also contains variations in scale and illumination. The proposed method covers the target till the last instance while IVT and LT significantly drifted from the target object. SM, DM and DFT covers extensively with variations at some time instances. From these results, it is observed that the proposed method achieves better tracking results in presence of scale and illumination variation. Finally, women sequence contains scale variations and abrupt motion. The proposed produces good performance in tracking but DFT tracks the object effectively than proposed method. From these extensive simulation results, it is clear that the Tree based model is significantly better than the 5 existing models namely IVT, LT, DFT, DM and SM respectively.

Table 2 Center Localization Error

Sequence	Various models					
	IVT	CT	DFT	Shallow	Deep	Tree-Structured
Boy	91.3	9	106	4.1	2.6	2.1
Car4	2.1	86	61.9	2.9	3.3	3.6
Football	14.3	11.9	9.3	5.4	11.2	3.2
Freeman1	11.6	119	10.4	7.6	11.2	6.7
Shaking	85.7	80	26.3	10.6	16.6	12.4
Woman	177	114	8.5	5.5	9.5	3.5
Walking	1.6	6.9	5.9	7.5	7.6	2.2

Table 3 Average overlap rate

Sequence	Various models					
	IVT	CT	DFT	Shallow model (SM)	Deep model (DM)	Tree-Structured
Boy	0.26	0.59	0.4	0.73	0.82	0.84
Car4	0.88	0.21	0.24	0.86	0.85	0.85
Football	0.56	0.61	0.66	0.70	0.47	0.78
Freeman1	0.43	0.14	0.39	0.51	0.54	0.56
Shaking	0.03	0.1	0.64	0.61	0.57	0.67
Woman	0.15	0.13	0.76	0.44	0.42	0.71
Walking	0.73	0.52	0.56	0.54	0.63	0.74

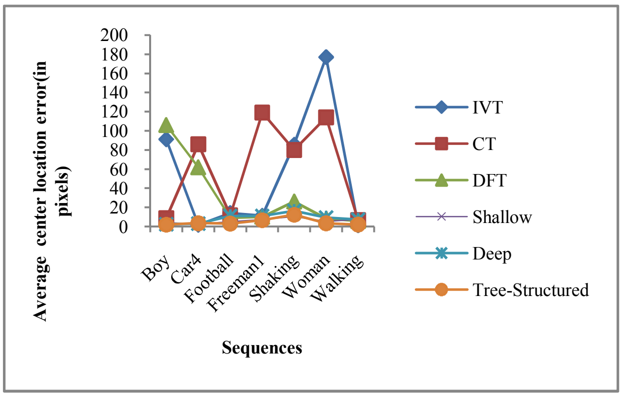

Fig 4 Comparison of the Average center location error (in pixels) for 7 benchmarks

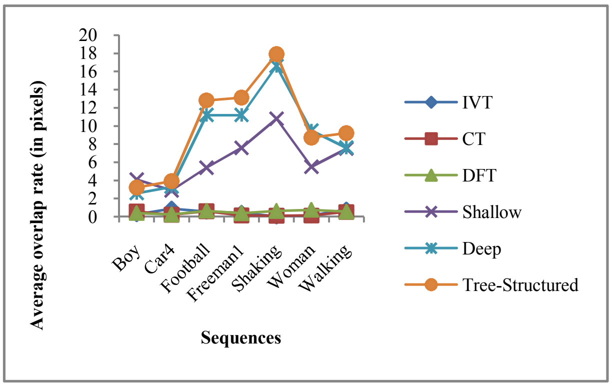

Fig 5 Comparison of the overlap rate for 7 benchmarks

CONCLUSION

In real scenario, it is very difficult to track the target object in complex situations (presence of fast motion, occlusion, etc.). To overcome these issues, Tree-based appearance modelis proposed with Bayesian framework. The proposed method characterizes the target into two levels: local level and global level. In local level, a group of local patches are utilized to map the target to adjust partial occlusion and scale variation. Reliability and stability indices are used to measure the reliability and stability of the tracked object. In the global level, the target is mapped by double bounding boxes based on foreground and background. The interior box includes the target region and the exterior box includes both the target and background region around the target. The global level accurately tracks the object in presence of motion blur and cluttered background. The performance of the proposed method is evaluated and compared with IVT, LT, DFT, DM and SM. The simulation results shows that the proposed method produces better results than existing models in terms of efficiency, accuracy and robustness.

References

1. G. Underwood, Cognitive processes in eye guidance: algorithms for attention in image processing, Cognit. Comput. (1) (2009) 64-76.
2. Y. Wu, J. Lim, M.-H. Yang, Online object tracking: a benchmark, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, Oregon, USA, 2013, pp. 2411-2418.
3. Pérez, C. Hue, J. Vermaak, and *et al.*, Color-based probabilistic tracking, in: Proceedings of the IEEE European Conference Computer Vision, Copenhagen, Danmark, 2002, pp. 661-675.
4. B. Babenko, M.-H. Yang, S. Belongie, Robust object tracking with online multiple instance learning, *IEEE Trans. Patt. Anal. Mach. Intell*. 33 (8) (2011) 1619-1632 .
5. X. Li, W. Hu, C. Shen, Z. Zhang, A. Dick, A.V.D. Hengel , A survey of appearance models in visual object tracking, *ACM Trans. Intell. Systems Technol*. 4 (4) (2013) 58.
6. R.T. Collins, Y. Liu, M. Leordeanu, Online selection of discriminative tracking features, IEEE Trans. Pattern *Anal. Mach. Intell*. 27 (10) (2005) 1631-1643.
7. D.A. Ross, J. Lim, R.S. Lin, *et al.*, Incremental learning for robust visual tracking, *Int. J. Comput. Vis.* 77 (1-3) (2008) 125–141.
8. X. Mei, H. Ling, Robust visual tracking using l1 minimization, in: Proceedings of the IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 2009, pp. 1436-1443
9. D.A. Ross, J. Lim, R.-S. Lin, M.-H. Yang, Incremental learning for robust visual tracking, *Int. J. Comput. Vis.* 77 (1–3) (2008) 125-141.
10. X. Jia, H. Lu, M.-H. Yang, Visual tracking via adaptive structural local sparse appearance model, in: CVPR, 2012, pp. 1822-1829.
11. D. Wang, H. Lu, M.-H. Yang, Online object tracking with sparse prototypes, IEEE Trans. Image Process. 22 (1) (2013).
12. K. Zhang, L. Zhang, M.-H. Yang, Real-time compressive tracking, in: ECCV, 2012, pp. 864-877.
13. S. Hare, A. Saffari, P.H. Torr, Struck: structured output tracking with kernels, in: ICCV, 2011, pp. 263-270.
14. N. Wang, D.-Y. Yeung, Learning a deep compact image representation for visual tracking, in: Advances in Neural Information Processing Systems, 2013, pp. 809-817.
15. R.T. Collins, Y. Liu, M. Leordeanu, Online selection of discriminative tracking features, IEEE Trans. *Pattern Anal. Mach. Intell*. 27 (10) (2005) 1631-1643.
16. M. Godec, P.M. Roth, H. Bischof, Hough-based tracking of non-rigid objects, Comput. Vis. Image Underst. 117 (10) (2013) 1245-1256.

17. D. Comaniciu, V. Ramesh, P. Meer, Kernel-based object tracking, IEEE Trans. *Pattern Anal. Mach. Intell.* 25 (5) (2003) 564-577.
18. B. Han, L. Davis, On-line density-based appearance modeling for object tracking, in: Proceedings of the Tenth IEEE International Conference on Computer Vision, Beijing, China, 2005, pp. 1492-1499.
19. H. Grabner, H, Bischof, On-line boosting and vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, 2006, pp. 260-267.
20. J.F. Henriques, R. Caseiro, P. Martins, *et al*., Exploiting the circulant structure of tracking-by-detection with kernels, in: Proceedings of the European Conference on Computer Vision, Florence, Tuscany, Italy, 2012, pp. 702-715.
21. Y. Wu, J. Lim, M.–H. Yang, Online object tracking: a benchmark, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, Oregon, USA, 2013, pp. 2411-2418.
22. V. Mahadevan, N. Vasconcelos, Biologically inspired object tracking using center surround saliency mechanisms, IEEE Trans. *Pattern Anal. Mach. Intell.* 35 (3) (2013) 541-554.
23. Z. Cai, L. Wen, J. Yang, *et al*., Structured visual tracking with dynamic graph, in: Proceedings of the Tenth IEEE Conference on Asian Conference on Computer Vision, Daejeon, Korea, Springer Berlin Heidelberg, 2013, pp. 86-97.
24. R. Yao, Q. Shi, C. Shen, *et al*., Part-based visual tracking with online latent structural learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, Oregon, USA, 2013, pp. 2363-2370.
25. M. Tang, X. Peng, Robust tracking with discriminative ranking lists, IEEE Trans. Image Process. 21 (7) (2012) 3273-3281.
26. J. Kwon, K.M. Lee, Highly nonrigid object tracking via patch-based dynamic appearance modeling, IEEE Trans. *Pattern Anal. Mach. Intell*. 35 (10) (2013), 2427-2441.
27. . Yao, Q. Shi, C. Shen, *et al*., Part-based visual tracking with online latent structural learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, Oregon, USA, 2013, pp. 2363-2370.
28. M. Tang, X. Peng, Robust tracking with discriminative ranking lists, IEEE Trans. Image Process. 21 (7) (2012) 3273–3281.
29. D.Y. Lee, J.Y. Sim, C.S. Kim, Visual tracking using pertinent patch selection and masking, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 3486–3493.
30. J.F. Henriques, R. Caseiro, P. Martins, J. Batista, High speed tracking with kernelized correlation filters, IEEE Trans. *Pattern Anal. Mach. Intell*. 37 (3) (2015) 583-596.
31. L. Cehovin, M. Kristan, A. Leonardis, Robust visual tracking using an adaptive coupled-layer visual model, IEEE Trans. *Pattern Anal. Mach. Intell*. 35 (4) (2013) 941-953.
32. D. Chen, Z. Yuan, Y. Wu, *et al*., Constructing adaptive complex cells for robust visual tracking, in: Proceedings of the IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 2013, pp. 1113-1120.
33. Bohan Zhuang, Lijun Wang, Huchuan Lu, Visual tracking via shallow and deep collaborative model, Neurocomputing 218 (2016) 61–71

How to cite this article:

Bhuvaneswari R and Ravi Subban (2018) 'Tree-Based Appearance Model for Robust Visual Tracking', *International Journal of Current Advanced Research,* 07(5), pp. 12722-12728. DOI: http://dx.doi.org/10.24327/ijcar.2018.12728.2246

****