2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

An Algorithm for Real-time Object Tracking in Complex Environment

Dongxu Gao, Jiangtao Cao and Zhaojie Ju

Abstract— **The current sparse representation tracking algorithm is not suitable for the objects that illumination changes, scale changes, the object color is similar with the surrounding region,and occlusion etc, what's more, it is hard to realize real-time tracking for solving an $l1$ norm related minimization problems. An optimal algorithm is introduced by exploiting an accelerated proximal gradient approach which contains some improvements of particle filter function, sparse representation alterative weights and coefficient. These improvements not only reduce the influences of appearance change but also make the tracker runs in real time. Both qualitative and quantitative evaluations demonstrate that the proposed tracking algorithm has favorably better performance than several state-of-the-art trackers using challenging benchmark image sequences, and significantly reduces the computing cost.**

I. INTRODUCTION

Visual tracking has long been an important research topic in the computer vision field as it is widely applied in the automated surveillance, vehicle navigation, automatic object identification and tracking the target. Lots of methods have have been successfully applied to object tracking and can be summarized into two categories: discriminative and generative approaches[1]. Discriminative methods treat the tracking problem as a classification problem, which aims to segment the target from the background [2][3]. So it considers the information of both the target and background. Some trackers combined a set of weak classifiers into a strong one [4], adopting an online boosting method to update discriminative features[5] or learning a large number of positive and negative samples for tracking. Generative methods formulated tracking by establishing the appearance model of the target. In order to develop effective models, several particularly factors should be considered. At the first, to adapt to the target appearance variations caused by pose change and illumination change, the target model needs to be updated online [6][7]. In addition, Yi Wu et al. [8] presented a novel Blur-driven tracker framework for tracking motion-blurred targets. Experimental results showed that discriminative models perform better when the training set size is large while generative models achieve higher generalization if limited data is available [9][10][11].

Dongxu Gao and Jiangtao Cao are with the Department of Information and Control Engineering, Liaoning Shihua University, Fushun, Liaoning, China (email: gdxaaa@163.com).

Zhaojie ju are with the Intelligent Systems and Biomedical Robotics group, School of Creative Technologies, University of Portsmouth, Portsmouth, England PO1 2DJ, UK(email: zhaojie.ju@port.ac.uk).

This work was supported by the National Natural Science Foundation of China under Project Code 61203021 and 61103123

Benefitting from less cost of generative model, sparse representation techniques [12][13], also known as compressed sensing or compressive sampling, for finding a sparse solution through solving an $l1$-regularized least problem, plays a critical role in both mathematics and applications such as visual tracking. Mei et al. [12][14][15] employed the sparse representation of the object as the appearance model and then obtained the tracking target by solving the $l1$ minimization problem, and achieved some ideal results but failed when there were similar objects in the scene. To further improve the method, they introduced non-negativity constrains and dynamically updated the target templates [16][17]. However, most sparse representation used the holistic model to denote the appearance model and cannot handle distracters and partial occlusion accurately.

Similarly with sparse representation, but combined with classic principal component analysis (PCA) algorithm and online object tracking algorithm, an robust generative tracking algorithm with an adaptive appearance model is proposed called sparse prototypes(SP), which handles partial occlusion built into the same framework of the online object tracker [18]. With the intuition that the appearance of a tracked target can be sparsely represented, therefore, finding a sparse which approximate the template subspace is necessary because it is effective in dealing with pose, illumination and scale variation as well as appearance change. Moreover, this paper also aims at developing a more efficient tracker in real time. The main contribution of this paper is to introduce a very fast numerical method [19] to solve the minimization problem [18], which is further reorganized, so that the proposed numerical method could be combined into the SP tracker and make it run in real time.

II. RELATED WORK

In recent years, there has been a large amount of literature research on target tracking problems, and studies which related to our work are summarized in this section. Generally speaking, there are several major issues in target tracking, such as appearance caused by in-plane rotation, scale illumination, poses change, partial occlusion and so on. Several experimental results demonstrate that PCA subspace representation with online update is effective in dealing with some of this issues expect partial occlusion [20]. The Incremental Visual Tracking(IVT) [21] method introduced an online update approach for efficiently learning and updating a low dimensional PCA subspace representation of the target object, which is sensitive to partial occlusion.

978-1-4799-1484-5/14/$31.00 ©2014 IEEE

However, there might be a drift by using the method of direct template update, so Viola et al.[22] introduced Multiple Instance Learning(MIL) into visual tracking to address this problem. Moreover, the $l1$ tracker [15] included a sparse representation of trivial templates so that its sparse linear combination can presents the occlusions and image noise in the target. In this approach, the sparse representation is obtained via solving a $l1$-norm related minimization problem and the $l1$ tracker turns out to be too slow to be a real time tracking method.

In this paper, advantages of both subspace method and s-parse representation method are combined effectively to solve the partial occlusion problem. In addition, a fast numerical method for solving the $l1$-norm related minimization problem is also applied to improve the method's computational rate. The experimental results show the proposed method has a good robustness against pose changes and illumination changes while at the same time achieving a great running time efficiency.

III. Introduction to sparse prototypes

This tracking method is closely related to the tracking method proposed by Xue Mei and Dong Wang[15][18], The main difference is the model representation and minimal method, in this paper, we can quickly get the result of the operation, we first give a brief review on the l1 tracker within the particle framework[18][19][21] for tracking the target.

Particle filtering [21] is to find a set of transmissions in the state space representation of a random sample to approximate the probability density function, instead of using the sample mean calculus, and then get the system state minimum variance estimation process, these samples are called as "particles" vividly, and therefore called particle filter. The representation is then used in the particle filter framework [18] for object tracking. Specially, for frame at t, we set the state variable x_t which indicates the position and shape information, with $z_{1:t}$ to describe the observed value of the target from the first frame to the frame t. The state prediction equation and status update equations for the particle filter are as follows:

$$p(x_t|z_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|z_{1:t-1})dx_{t-1} \tag{1}$$

$$p(x_t|z_{1:t}) = \frac{p(z_t|x_t)p(x_t|z_{1:t-1})}{p(z_t|z_{1:t-1})} \tag{2}$$

Where $p(z_t|x_t)$ denotes the observation likelihood from the state at time t. It is practically intractable to calculated the above probability distribution directly, so the posterior $p(x_t|z_{1:t})$ is approximated by a finite samples $\{x_t^1, x_t^2, \cdots, x_t^N\}$ with different weights $\{w_t^1, w_t^2, \cdots, w_t^N\}$ where N is the number of samples and $\delta(x_t - x_t^i)$ is Dirac delta function. The samples are generated by an approximated equation

$$p(x_t|z_{1:t}) \approx \sum_{i=1}^{N} w_t^i \delta(x_t - x_t^i) \tag{3}$$

And the weights are updated by:

$$w_k^i = w_k^{i-1} \frac{p(z_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{t-1}^i, z_t)} \tag{4}$$

In the case of the bootstrap filter $q(x_t^i|x_{t-1}^i, z_t) = p(x_t|x_{t-1})$ and the weights become the observation likelihood $p(z_t|x_t)$.

According to the weights distribution, in each step, samples are re-sampled to generate new sample set with equal weights in case the weights of some particles keep increasing or fall into the degeneracy case.

The sparse prototypes aims at calculating the observation likelihood $p(z_t|x_t)$, and some researchers have proposed an algorithm [12][16][17] by casting the tracking problem as finding the optimal patch with sparse representation and handling partial occlusion with trivial templates then the patch is normalized and reshaped to a one-dimensional vector y which is formulated as a target candidate. This can be viewed as a minimum error reconstruction through a regularized $l1$ minimization function with non-negativity constraints:

$$min_c \frac{1}{2} \|y - Bc\|_2^2 + \lambda \|c\|_1 \tag{5}$$

Where B is composed of target template set and trivial template sets, while c is composed of target coefficient and trivial coefficient.

The main differences compared with sparse representation lie in a different target template model. For target tracking, we model object appearance with PCA basis vectors and some trivial templates. The sparse prototypes representation model is then:

$$y = Dz + e = \begin{bmatrix} D & I \end{bmatrix} \begin{bmatrix} z \\ e \end{bmatrix} \tag{6}$$

Where I is identity matrix and is a trivial template set, y indicates an observation vector, D represents a column basis vectors, z represents the coefficients of basis vectors, and e denotes the error term. The prototypes in this formulation consist of a set of PCA basis vectors and a set of trivial templates. We solve Eq. (6) by

$$min_a \frac{1}{2}||y - Ba||_2^2 + \lambda||a||_1 \tag{7}$$

Where $B = [D, I, -I]$ is composed of target template set D and trivial template sets I and $-I$. a represents the coefficient corresponds to B.

Finally, the observation likelihood is derived from (see [18] and the references therein):

$$p\left(z_t^i|x_t^i\right) = \frac{1}{\Gamma} \exp[-(||w^i \circ (y_t^i - Dz^i)||_2^2 + \beta \sum (1 - w^i))] \tag{8}$$

Which consider the occlusion kept. Where $p\left(z_t^i|x_t^i\right)$ indicates not only the reconstruction error of Eq.(7), but also any pixel as being occluded. z is obtained by solving the Eq.(7), α is a constant controlling the shape of the Gaussian kernel, Γ is a normalization factor, then the optimal state x_t^* of frame t obtained by:

$$x_t^* = \arg\max_{x_t^i \in S_t} p\left(z_t|x_t^i\right) \tag{9}$$

In addition, the update of observation model is adopted [18] for handling appearance change of a target object for visual tracking. Therefore, the tracking result is chosen from the candidate of the maximum observation likelihood when tracking at time t.

IV. Real time object tracking

Through the narrative of the front sections, the SP algorithm can be seen as an optimization problem by solving the Eq.(7). The proposed method for solving the $l1$ minimization Eq. (7) based on the accelerated proximal gradient approach[19][24] is given in algorithm 1.

Algorithm 1 our tracker

1. Initialization
2. Load the image sequences
3. Candidate Sampling
4. Adopt $PCA\_ l1$ to acquire templates set
5. Solving the minimization Eq.(7) via algorithm 2
6. Calculate observation likelihood via Eq.(8)
7. Obtain the optimal candidate by Eq.(9)
8. Collect samples for model update

The minimization model can be converted to the APG method:

$$\min F(a) + G(a) \tag{10}$$

As long as $F(a)$is a convex function with Lipschitz [19] continuous gradient, as the same time, $G(a)$is a non-smooth but convex function. Within $K = O(\sqrt{L/\epsilon})$ iterations, x_k achieves ϵ-optimality such that $\|x_k - x^*\| < \epsilon\|$, where x^* is one minimizer of Eq.(10).

For Eq.(7),it is completely equivalent to the minimization problem as below:

$$argmin_a \frac{1}{2}||y - Ba||_2^2 + \lambda \mathbf{1}_T^T a + ||I||_1 + \mathbf{1}_{R_+^n}(-I) \tag{11}$$

in which $\mathbf{1} \in \mathrm{R^n}$ represent the vector with all entries are equal to 1, and $\mathbf{1}_{\mathrm{R_+^n}}(\mathrm{a})$ can be defined:

$$\mathbf{1}_{\mathrm{R_+^n}}(\mathrm{a}) = \begin{cases} 0 & \mathrm{a} \geq 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{12}$$

Then we assume that:

$$\begin{array}{c} F(a) = \frac{1}{2}||y - Ba||_2^2 + \lambda \mathbf{1}_T^T a \\ G(a) = ||I||_1 + \mathbf{1}_{R_+^n}(-I) \end{array} \tag{13}$$

If the result of Eq.(11) is x_{k+1}, then the minimization problem for Eq.(11) has the solution which is given in algorithm 2.

Algorithm 2 Real time numerical algorithm for solving the minimization Eq.(7)

1. For a given nonnegative vector λ, choose $x^0 = x^1 = 0 \in \Re^n, t^0 = t^{-1} = 1$.
   For $k = 0, 1, 2, \cdots$, generate x^{k+1} from x^k according to the following iteration which would be convergence eventually:
2. $y_k := x_k + \frac{t_{k-1}-1}{t_k}(x_k - x_{k-1})$
3. $g_k := y_k - \nabla f(y_k)/L,$
4. $x_{k+1} := max(0, g_k)$
5. $t_{k+1} := \frac{1+\sqrt{1+4(t_k)^2}}{2}$

V. Experiments

Our algorithm is implemented in Matlab and achieves about average 23 frames per second with 400 particles on a PC with Intel E7500 CPU (2.93GHz).

A. Experimental settings

Target of interest in the first frame position will be set manually and then automatic tracking is possible. The object is normalized to 32×32 pixels and 16 eigenvectors are used in all experiments for PCA representation. The number of templates for the sparse representation is 10. 600 particles are used and the maximum number of iterations is set to 20 in algorithm 2. The tracker is updated every 5 frames and the regularization constant λ is assumed as 0.05 in all experiments.

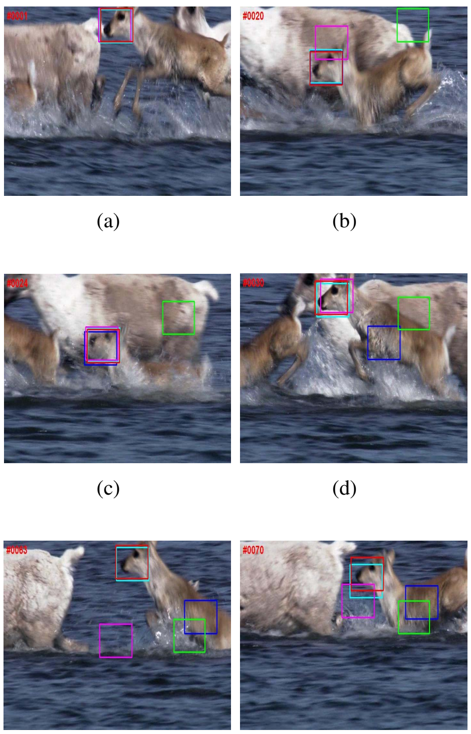
B. Qualitative comparison with other methods

In order to illustrate the qualitative comparison more clearly, some methods are described briefly here.The Visual Tracking Decomposition(VTD) method [25] used the observation model which is decomposed into multiple basic observation models that are constructed by sparse principal component analysis (SPCA) of a set of feature templates.The MIL method [22] put all ambiguous positive and negative samples into bags to learn a discriminative model for tracking.The $l1$ method [15] adopted the holistic representation of the object as the appearance model and then tracks the object by solving the $l1$ minimization problem.The assessment of several methods above in different situations are shown as below:

1) Heavy occlusion:

Fig.1 represents the identification of different methods. Fig.2 illustrates the tracking results from seven challenging sequences with significant change of scale, illumination and pose variation, as well as occlusion. For the occlusion sequence, there is a serious occlusion, while the caviar1 sequence's target is occluded by a similar object. As shown in Fig.2, our algorithm and $l1$ methods have better performance, since both methods take occlusion into consideration and handle occlusion using sparse representation with trivial templates while others can not deal with appearance changes caused by the pose and occlusion. Although the VTD method is able to track the object, it can not calculate the in-plane rotation because of the design of affine motion model. On the other hand, the deviations of $l1$ tracker is unacceptable especially when partial occlusion occurs (e.g.,#559) . This may be caused by the fact that the occlusion is ignored when $l1$ tracker makes new image observation.

Fig. 1. The identification of different methods



(a) (b)

(c) (d)

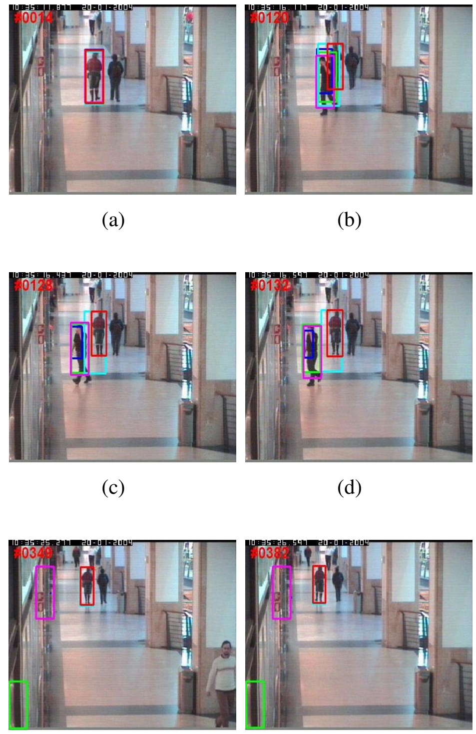
(e) (f)

Fig. 2. The sequences of Deer

In Fig.3, the caviar1 sequence is challenging as it contains similar objects, but our method performs best especially for targets occluded by similar objects. The PCA, $l1$ method and MIL trackers, however, drift away from such targets. The VTD method does not perform well but it can track the object as the generalized features are used for object representation.

2) Illumination change:

For the car4 sequence, as shown in Fig.4 and Fig.5, there will be dramatic illumination changes when the vehicle passes through the shadows of the trees or the overpass. The MIL methods are less effective when tracking the car, and the tracking results even drift away from the target because of the illumination variation. Both our methods and PCA methods can track the target effectively while others have serious drift phenomenon. This may be caused by the changes in appearance of the object, which can be well approximated with fixed posture subspace. Additionally, from the above results, trackers, such as MIL tracker, $l1$ tracker and VTD



(a) (b)

(c) (d)

(e) (f)

Fig. 3. The sequences of Caviar1

tracker do not adapt to different scales or in-plane rotations especially dramatic lighting changes.

3) Cluttered background:

When it comes to lemming sequence, the most prominent and challenging issues is the heavy occlusion in a cluttered background as well as changes of scales and poses as shown in Fig.6 and Fig.7. The MIL and our methods perform well all the time when the abrupt motion in a cluttered or in-plane rotation occurs and the surrounding region has similar textures. The VTD method can track the target accurately even after drifting away, while the 1 method cannot since the target object is similar to the surrounding background. The PCA method fails to track the whole object and only tracks the center of the object after the abrupt motion occurs due to the different appearance update strategies.

4) Abrupt motion:

In the Deer video, there is a huge challenge about the varying appearance caused by the motion blur as well as the fast motion. The tracking results are shown in Fig.8 and Fig.9. For the deer sequence, our tracker and VTD tracker perform better than other methods due to the re-initialization update model. The MIL tracker is able to track the object in some frames after drifting away caused by the similar surrounding background as the deer reappears at the same location in the image. But it failed later due to the motion blur. The PCA and $l1$ trackers are less effective and drift away from the deer with the repetitive motions.

C. Quantitative comparison with other methods

Table I and Table II summarizes the results in terms of the

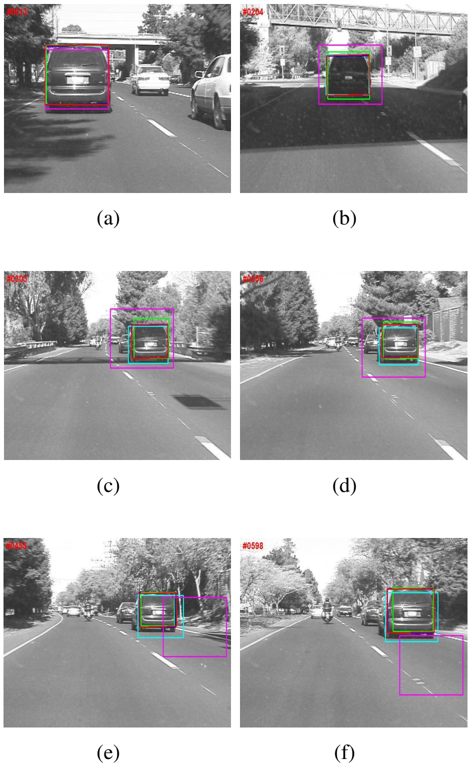


Fig. 4. The sequences of car4

Fig. 5. The sequences of DavidIndoor

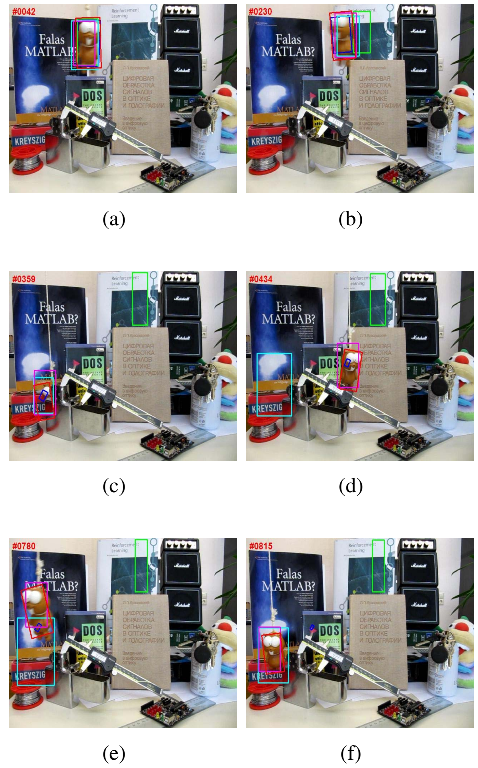


Fig. 6. The sequences of lemming

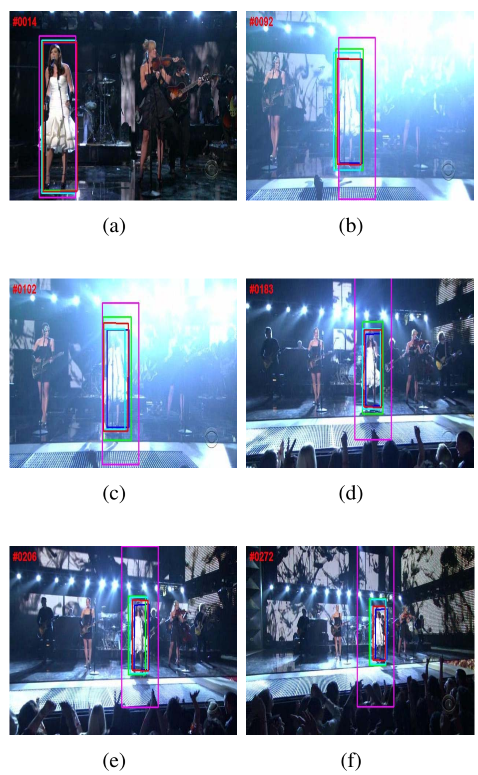
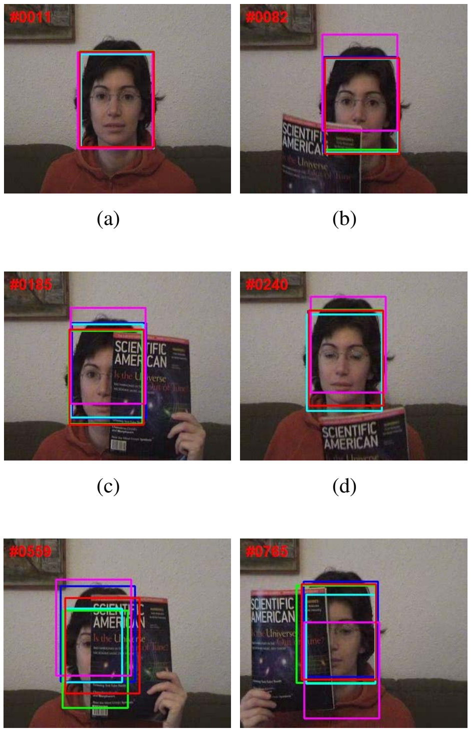


Fig. 7. The sequences of Singer1



(a) (b) (c) (d) (e) (f)

Fig. 8. The sequences of Occlusion1

(a) (b) (c) (d) (e) (f)

Fig. 9. The sequences of car11

average tracking overlap and error, which shows our method achieves the lowest tracking errors. In addition, The success rate is evaluated through the overlap rate which is defined by the PASCAL VOC [26] criterion if given the tracking result of each frame R_T, the corresponding ground truth R_G and $score = \frac{area(R_T \cap R_G)}{area(R_T \cup R_G)}$. The tracking results are regarded as being valid when the score is over 0.5. The average overlap rate of our tracker is 0.82 while the highest is 0.61 at present.

TABLE I
OVERLAP RATE OF TRACKING METHOD

Overlap Rate	PCA	L1	VTD	MIL	Ours
Car4	0.92	0.84	0.73	0.34	0.92
Car11	0.81	0.44	0.43	0.17	0.81
Caviar1	0.28	0.28	0.83	0.25	0.89
David Indoor	0.71	0.63	0.53	0.45	0.80
Lemming	0.18	0.13	0.35	0.53	0.75
Occlusion	0.85	0.88	0.77	0.59	0.91
Deer	0.22	0.04	0.58	0.21	0.61
Singer1	0.66	0.70	0.79	0.33	0.82
Average	0.58	0.49	0.63	0.36	0.81

TABLE II
AVERAGE CENTER ERROR

Average center	PCA	L1	VTD	MIL	Ours
Car4	2.87	4.08	12.29	60.10	3.03
Car11	2.11	33.25	27.05	43.47	2.17
Caviar1	45.25	119.93	3.91	48.50	1.67
David Indoor	3.59	7.63	13.55	16.15	3.65
Lemming	93.38	184.85	86.89	25.58	9.1457
Occlusion1	9.18	6.50	11.13	32.26	4.70
Deer	127.47	171.47	11.92	66.46	8.53
Singer1	8.48	4.57	4.06	15.17	4.75
Average	36.54	66.54	21.35	38.46	4.71

Consistent with the appearance updating in [18] , if the object is well tracked and the occlusion rate is small, the tracking result is then used to update the observation model directly. We give the results using only Eq.(8) with the occlusion map in Table I to demonstrate how the occlusion map facilities the object tracking and the observation update. The results show that our algorithm can effectively predict the occlusion maps and further improve the tracking results in terms of both overlap rate in Table1 and the center location error in Table II. Overall, the minimum error rate and maximum overlap rate in all the sequences show that our algorithm has the best performance compared with several other the-state-of-the-art trackers on challenging benchmark image sequences by considering more factors.

D. Computational Complexity

Performance evaluation methods above show the proposed method is more accurate and robust in most challenging sequences. In addition, it is more efficient, because it reduces several orders of magnitude for the computational complexity compared with the other algorithms. We have analysed the time complexities of some algorithms previously mentioned, for example, the complexity of the $l1$ tracker is (O$(d^2 + dk)$), the sparse representation tracker has (O(ndk)) complexity

[18] and (O(dk)) for IVT which is more faster compared with the others, but for the SP tracker in this paper, the complexity is O($\sqrt{d/k}$) , which is consistent with the analysis of accelerated proximal gradient algorithm [19], and the average time for solving one image patch(32×32) is 0.039ms, by contrast, the average time for IVT is 0.19ms.

VI. Conclusion

The method in this paper exploits both classic principal component analysis algorithm and recent sparse representation to obtain the observation model and uses an accelerated proximal gradient approach to locate the target more accurately and faster. It does not only adapt the tracker to account for the object appearance change but also handles the occlusion caused by similar objects or distinct. Comparing with several other the-state-of-the-art methods,the experimental results on challenging benchmark image sequences demonstrated both the effectiveness and efficiency of the proposed method. As the proposed algorithm utilized representation scheme and the minimum problem, it can still be further optimized.

References

[1] J. Gallego, M. Pardas and G. Haro, "Enhanced foreground segmentation and tracking combining Bayesian background,shadow and foreground modeling," *Pattern Recognition Letters* vol.33,pp.1558-1568, 2012.

[2] M. Heber, M. Godec, M. Ruther, P.M. Roth, H.Bischof, "Segmentation based tracking by support fusion," *Computer Vision and Image Understanding* vol. 117, pp.573-586, 2013.

[3] W. Brendel and S. Todorovic, "Video Object Segmentation by Tracking Regions,"*IEEE Int. Conf. on Computer Vision (ICCV)*,pp. 833–840, 2009.

[4] S.Avidan, "Ensemble tracking," *IEEE Transaction on pattern analysis and machine intelligence*, vol. 29, no. 2,pp. 261-271, 2007.

[5] H.Grabner and H.Bischof. "On-line boosting and vision," *Computer Vision and Pattern Recognition (CVPR)*, pp. 260-267,2006.

[6] Y. Wu, J. Lim and M. H. Yang, "Online Object Tracking: A Benchmark,"*Computer Vision and Pattern Recognition (CVPR)*,pp. 2411-2418, 2013.

[7] R. Yao, Q. Shi, C. Shen, Y. Zhang and A. Hengel, "Part-based Visual Tracking with Online Latent Structural Learning," *Computer Vision and Pattern Recognition (CVPR)*, pp. 2363-2370, 2013.

[8] Y. Wu, H. Ling, J. Yu, F. Li, X. Mei and E. Cheng, "Blurred Target Tracking by Blur-driven Tracker," *Computer Vision (ICCV)*, pp. 1100-1107, 2011.

[9] J. A. Lasserre, C. M. Bishop, and T. P. Minka, "Principled hybrids of generative and discriminative models,"*Computer Vision Pattern Recognition (CVPR)*, pp. 87-94, 2006.

[10] A. Y. Ng and M. I. Jordan, "On discriminative versus generative classifiers: A comparison of logistic regression and naive bayes," *in Proc. Adv. Neural Inform. Process. Syst.*, pp. 438-451, 2002.

[11] W. Zhong, H. Lu, and M. H. Yang, "Robust object tracking via sparsity based collaborative model,"*Computer Vision and Pattern Recognition (CVPR)*, pp. 1838-1845, 2012.

[12] B. Liu, J. Huang, L. Yang, and C. A. Kulikowski. "Robust tracking using local sarse appearance model and k-selection," *Computer Vision and Pattern Recognition (CVPR)*, pp. 1313-1320, 2011.

[13] B. Liu, L. Yang, J. Huang, P. Meer, L. Gong, and C. A. Kulikowski, "Robust and fast collaborative tracking with two stage sparse optimization," *European Conference on Computer Vision (ECCV)*, pp. 624-637, 2010.

[14] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 210-227, 2009.

[15] X. Mei and H. Ling, "Robust visual tracking using $l1$ minimization,"*Computer Vision (ICCV)*, pp. 1436-1443,2009.

[16] X. Mei, H. Ling,Y.Wu,E.Blasch, and L.Bai, "Minimum error bounded efficient $l1$ tracker with occlusion detection," *Computer Vision and Pattern Recognition (CVPR)*, pp. 1257-1264, 2011.

[17] X. Mei and H. Ling. "Robust Visual tracking and vehicle classification via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* vol. 33, no. 11, pp. 2259-2272, 2011.

[18] D. Wang, H. Lu and M. H. Yang. "Online Object Tracking With Sparse Prototypes,"*IEEE transactions on image processing*,vol. 22, No. 1, pp. 314-325, 2013.

[19] Z. Shen, K. C. Toh, S. Yun, "An Accelerated Proximal Gradient Algorithm for Frame-Based Image Restoration via the Balanced Approach," *SIAM J. Imaging Sciences*,pp.573-596, 2011.

[20] D. Wang, H. Lu, and Y. W. Chen, "Incremental MPCA for color object tracking,"*International Conference on Pattern Recognition (ICPR)*, pp. 1751-1754, 2010.

[21] D.Ross, J.Lim, R.S. Lin, and M.H. Yang, "Incremental Learning for robust visual tracking,"*International Journal of Computer Vision (IJCV)*, pp. 125-141, 2008.

[22] P.A. Viola, J.C. Platt, and C.Zhang, "Multiple instance boosting for object detection," *In Advances in Neural Information Processing Systems*, vol. 18, pp. 1417-1424, 2006.

[23] C.Huiying and L.Youfu, "Optimized particles for 3D trackers," *International Journal of Humanoid Robotics*, vol. 08, pp. 749-761, 2011.

[24] W.Zuo, Z.Lin, "A generalized accelerated proximal gradient approach for total-variation-based image restoration," *IEEE transactions on image processing*,vol. 20, no.10, pp. 2478-2759, 2011.

[25] J. Kwon, K. M. Lee, "Visual Tracking Decomposition,"*Computer Vision and Pattern Recognition (CVPR)*, pp. 1269-1276, 2010.

[26] M. Everingham, L. V. Gool, C. K. I. Williams, J. Winn and A. Zisserman, "The pascal visual object classes(voc) challenge,"*International Journal of Computer Vision(IJCV)*, pp. 303-338, 2010.