# CONTEXT MULTI-TASK VISUAL OBJECT TRACKING VIA GUIDED FILTER

*Yong Wang*, Xinbin Luo†, Shiqiang Hu‡*

*School of Electrical and Computer Engineering, University of Ottawa.
†School of Electronic information and Electrical Engineering, Shanghai Jiao Tong University
‡School of Aeronautics and Astronautics, Shanghai Jiao Tong University

## ABSTRACT

In this paper, we formulate particle filter based tracking as a multi-task sparse learning problem that exploits context information. The target and context information which modeled as linear combinations of principal component analysis (PCA) basis is formed as dictionary templates. We treat the dictionary templates as the guidance and the incoming candidates are filtered depending on the similarity between the guidance image and each input. The guided filter can help to distinguish the target from numerous candidates via context information. Then multi-task sparse learning is employed to learn the target and context information. The proposed learning problem is efficiently solved using an alternating direction method of multipliers (ADMM) method that yield a sequence of closed form updates. We test our tracker on challenging benchmark sequences that involve drastic illumination changes, large pose variations, and heavy occlusion. Experimental results show that our tracker consistently outperforms state-of-the-art trackers.

***Index Terms—*** Context information, multi-task sparse learning, guided filter

## 1. INTRODUCTION

Visual tracking is important for human computer interaction, video surveillance, action recognition to name a few. Although a lot of tracking algorithms have been proposed, it still remains a challenging problem for complex scenarios.

Context information has been applied actively in object detection [1], object classification [2], and object recognition [3]. Recently, it has been introduced into visual tracking [4, 5, 6]. The improved performance of these trackers is attributed to the use of context information in determining the target location. A new tracker has been proposed that combines the concept of context and structure for object tracking [4]. In [5], the authors exploit context information to guide tracking. In our point of view, if the target representation incorporates more background information, it in turn prevents the tracker to drift from the target into the background.

We model the tracking target with principal component analysis (PCA) basis vectors similar to [16]. Linear combination of the basis vectors can well represent uncorrupted samples. The guided filter has been introduced into object tracking [7] and showed good tracking results. The guided filter has edge-preserving smoothing property. It can be used for noise reduction, detail enhancement and so on. We treat each template as the guidance image. Thus, the intrinsic characteristics of the target are preserved and the illumination, background clutter effect are alleviated. With the aid of guided filter, the tracker can distinguish the target from the background through considering the appearance information between the guidance and the candidate images.

To accelerate the tracking results and improve tracking accuracy, multi-task sparse learning has been used in visual object tracking [8]. The decomposition form of multi-task sparse learning can be used to simultaneously capture a common set of features among relevant tasks and identifies outlier tasks [9, 10]. In this paper, the representation model is decomposed into a row-sparse matrix which corresponding to the overlapping features and an elementwise sparse matrix which corresponding to the non-shared features or outliers.

Inspired by the above work, we develop a computationally efficient, multi-task sparse learning tracker that exploits context information. The contributions of this paper are threefold.

1) We propose a multi-task sparse learning method for object tracking, which makes use of context information for more robust performance. To the best of our knowledge, this is the first work to exploit context information via multi-task sparse learning in object tracking.

2) The guided filter is used in subspace to alleviate the effect of illumination and background clutter. The guided filter preserves the edge information in the subspace. That is, most of the characteristics of the target can be retained. To the best of our knowledge, this is the first work to explore the guided filter in the subspace.

3) The decomposition model is employed to separate the common part and outliers in the candidates. This model enables our tracker to choose the appropriate candidate effec-

This work was jointly supported by the National Natural Science Foundation of China (No. 61374161) and China Aviation Science Foundation (No. 20142057006).

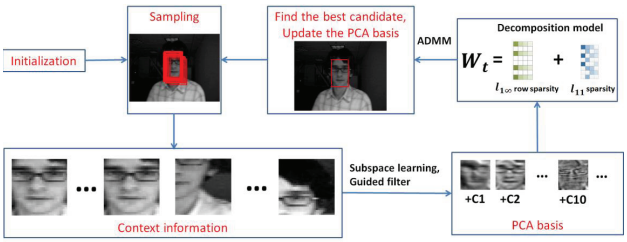


**Fig. 1**. The proposed tracking framework.

tively.

## 2. PROPOSED ALGORITHM

In this section, we give a detailed description of our particle filter-based tracking method, which makes use of context information and guided filter in a multi-task sparse learning framework to represent particle samples. Fig. 1 illustrates the overall flows of the proposed tracking framework.

### 2.1. Context Model

We assume an affine motion model between consecutive frames similar to [11, 5]. Thus, the state variable $ss_t$ consists of the six parameters of an affine transformation including of a 2-D linear transformation and a 2-D translation. The state transition distribution $p(ss_t|s_{t-1})$ is modeled to be Gaussian with the components of $sst$ assumed independent. The similarity between a target candidate (particle) and dictionary templates is defined as the observation model $p(y_t|ss_t)$. The model $p(y_t|ss_t)$ is computed as a function of the reconstruction error obtained by linearly representing $y_t$ using the target template dictionary. The particle that maximizes this function is selected to be the tracked target in each frame.

In the $t$th frame, n particles are randomly sampled around the previous states according to zero-mean Gaussian distributions. Each observation of the $i$th particle is represented as a linear combination of dictionary templates $D$. The dictionary $D$ is composed by two types of dictionary $D_T$ and $D_C$, which incorporate information about the target and context, respectively. The two dictionaries focus on complimentary aspects of particle representation: 1) target information; and 2) context information.

In each frame, the tracking target and all candidates with $k$ PCA basis vectors ($X_t$).

$$Y_t = D_t X_t, \tag{1}$$

where $Y_t \in R^{d*N}$, $Y_t = [y_1, y_2, \cdot, y_N]$ is the observation vector, $N$ is the number of obervations, $X_t \in R^{k*N}$ denotes a matrix of PCA basis vectors (d represents feature dimension and $k$ the number of PCA basis), $D_t \in R^{d*k}$ denotes the corresponding coding vectors (target coefficients).

### 2.2. Guided Filter

In this section, we give a brief introduction of guided filter. More detail can be found in [12]. The guidance image (subspace), filtering input image and filtering output image are denoted as $I$, $p$ and $q$ respectively. The guided filter can be represented by a local linear model:

$$q_i = a_j * I_i + b_j, \forall I \in \omega_j \tag{2}$$

where $i$ is the index of a pixel, and $j$ is the index of a local square window $w$ with a radius $r$. The minimization of the reconstruction error between $p$ and $q$ is as follows:

$$a_j = \frac{\frac{1}{|\omega|}\Sigma_{I\in\omega_j} I_i p_i - \mu_j \bar{p}_j}{\sigma_j^2 + \epsilon} \tag{3}$$

$$b_j = \bar{p}_j - a_j \mu_j \tag{4}$$

where $\mu_j$ and $\sigma_j$ are the mean and variance of $I$ in the window $j$, and $\epsilon$ is a regularization parameter controlling the degree of smoothness. The filtering output is computed by:

$$q_i = \bar{a}_i I_i + \bar{b}_i \tag{5}$$

where $\bar{a}$ and $\bar{b}$ are the average of $a$ and $b$ respectively on the window $w_i$ centered at $i$.

### 2.3. Decomposition Model

In our sampling scheme, most of the candidates share a common part of the target. Each of the candidates possesses its own characters that distinguish it from the others. The decomposition model [9, 10] in multi-task sparse learning can give an appropriate representation of this scenario. Therefore, the equation (1) can be re-written as the following form:

$$Y_t = D_t X_t = D_t(L_t + S_t) \tag{6}$$

The above equation can be transformed to the following equation:

$$\begin{aligned} min_{L_t,S_t} ||Y_t - D_t(L_t + S_t)||_F^2 + \lambda_1 ||L_t||_{1,\infty} \\ +\lambda_2 ||S_t||_{1,1}, X_t = L_t + S_t \end{aligned} \tag{7}$$

where $L_t$ is the row group sparsity component and $S_t$ is the elementwise sparsity component. The two parameters $\lambda_1$ and $\lambda_2$ balance reliable construction and joint sparsity, where $\lambda_1$ regulates the row group sparsity on $L_t$ and $\lambda_2$ controls the elementwise sparsity on $S_t$. The $l_{1,\infty}$ norm represent the common part of the candidates and the $l_{1,1}$ norm represent the individual property. In the ADMM algorithm [8, 13], the equation (7) is decomposed into two sub-problems (i.e., below equations (8) and (9)), and then each of the two sub-problems is iterated until convergence:

$$L_t : min_{L_t} \frac{1}{2} \parallel (Y_t - D_t S_t) - D_t L_t \parallel_F^2 + \lambda_1 \parallel L_t \parallel_{1,\infty} \tag{8}$$

Input: $Y_t$, $D_t$,

Initialize L, S, (here L and S represent $L_t$ and $S_t$, respectively; $t$ is omitted for clarity of the algorithm description in the following.)

k=1,

While stopping criterion is not met do

Solve the convex optimization problem:

$$L^{k+1} = \underset{L^k}{\text{minimize}} \frac{1}{2}\left\|\left(Y_t - D_t^k S^k\right) - D_t^k L^k\right\|_F^2 + \lambda_1 \left\|Z_L^k\right\|_{1,\infty}$$

$$s.t.\ L^k = Z_L^k,$$

Perform the iterations of scaled ADMM algorithm:

$$L^{k+1} = \left(\left(D_t^k\right)^T D_t^k + \rho I\right)^{-1}\left[\left(D_t^k\right)^T\left(Y_t - D_t^k S^k\right) + \rho\left(Z_L^k - u_L^k\right)\right],$$

$$Z_L^{k+1} = prox_{1inf}\left(Z_L^k\right),$$

$$u_L^{k+1} = u_L^k + L^{k+1} - Z_L^{k+1},$$

where $prox_{1inf}$ denotes proximal method for $\|*\|_{1,\infty}$ [14].

Solve the convex optimization problem with updated $L^{k+1}$:

$$S^{k+1} = \underset{S^k}{\text{minimize}} \frac{1}{2}\left\|\left(Y_t - D_t^k L^{k+1}\right) - D_t^k S^k\right\|_F^2 + \lambda_2 \left\|Z_S^k\right\|_{1,1}$$

$$s.t.\ S^k = Z_S^k,$$

Perform the iterations of scaled ADMM algorithm:

$$S^{k+1} = \left(\left(D_t^k\right)^T D_t^k + \rho I\right)^{-1}\left[\left(D_t^k\right)^T\left(X_t - D_t^k L^{k+1}\right) + \rho\left(Z_S^k - u_S^k\right)\right],$$

$$Z_S^{k+1} = prox_{11}\left(Z_S^k\right),$$

$$u_S^{k+1} = u_S^k + S^{k+1} - Z_S^{k+1},$$

where $prox_{11}$ denotes proximal method for $\|*\|_{1,1}$.

k = k+1;

end while

Output: solution $L_t = L^{k+1}$, $S_t = S^{k+1}$, $X_t = L_t + S_t$, to equation (7).

**Fig. 2**. Implementation of the proposed multi-task sparse learning algorithm using ADMM.

$$S_t : min_{S_t} \frac{1}{2} \parallel (Y_t - D_t L_t) - D_t S_t) \parallel_F^2 + \lambda_2 \parallel S_t \parallel_{1,1} \quad (9)$$

We summarize our proposed algorithm implemented by ADMM for resolving equation (7) in Fig. 2.

## 3. EXPERIMENT

Our tracker is implemented in MATLAB, which runs at 1 FPS on a 2.53 GHz CPU with 4GB RAM. We set the parameters fixed for all sequences. Two criteria are adopted in our experiments [15]: the center location error and the overlapping rate, which are calculated between the tracking target box $R_t$ and the ground truth rectangle box $R_g$. The center location error is computed by the average distance between the center of $R_t$ and $R_g$. The overlapping rate is calculated by area $\frac{area(R_t \bigcap R_g)}{area(R_t \bigcup R_g)}$.

We test the proposed algorithm and 23 state-of-the-art approaches on eight publicly available sequences which contain various challenges such as illumination changes, partial occlusion, pose variation and complex background, to name just a few. The eight trackers are the Incremental Visual Tracking (IVT) [16], L1 tracking (L1T) [11], L1-APG tracking [17], multi-task tracking (MTT) [8], Multiple Instance Learning tracking (MIL) [18], compressive tracking (CT) [19], L2-RLS [20] , WMIL [21].

| | CT | IVT | L1-APG | L1 | L2-RLS | MIL | MTT-L01 | WMIL | Our1 |
|---|---|---|---|---|---|---|---|---|---|
| basketball | 63.3142 | 163.4504 | 103.3235 | 407.3348 | 133.4781 | 99.1368 | 114.8406 | 28.47.7 | 26.514 |
| car | 3.8078 | 26.1703 | 14.637 | 33.3288 | 21.1479 | 6.315 | 3.303 | 96.9464 | 5.067 |
| car11 | 3.8078 | 2.0304 | 1.9286 | 33.3288 | 2.7366 | 7.6508 | 2.2616 | 96.9464 | 3.392 |
| caviar2 | 63.1670 | 13.3922 | 5.8703 | 3.4694 | 16.3851 | 22.6452 | 4.8253 | 62.0663 | 7.179 |
| davidIndoor | 16.5161 | 67.5226 | 31.2135 | 221.8834 | 20.8459 | 23.1548 | 88.3262 | 23.577 | 15.3078 |
| dog1 | 15.7522 | 45.5823 | 9.7048 | 260.6967 | 4.67428 | 16.3329 | 8.2277 | 29.521 | 11.856 |
| faceocc | 18.8312 | 64.9672 | 24.8161 | 111.9744 | 20.7714 | 32.9755 | 33.3148 | 49.0825 | 11.169 |
| faceocc2 | 24.5455 | 63.7754 | 12.4189 | 153.9712 | 11.5001 | 21.4552 | 8.1904 | 30.9168 | 8.741 |
| juice | 6.7165 | 69.1284 | 0.9835 | 110.9964 | 6.1545 | 42.3063 | 3.4975 | 10.4687 | 5.856 |
| shop | 70.5524 | 7.8512 | 2.9598 | 3.8836 | 61.0245 | 121.302 | 2.4224 | 62.8289 | 2.829 |
| singer | 19.2334 | 47.1161 | 5.098 | 386.1092 | 27.0821 | 23.1601 | 51.9938 | 18.0611 | 7.181 |
| ucsdpeds | 5.3104 | 11.4702 | 1.7455 | 101.5581 | 62.797 | 10.9856 | 1.2932 | 12.555 | 2.833 |
| david2 | 80.5788 | 48.9736 | 2.6076 | 126.654 | 1.965 | 56.5513 | 1.4139 | 11.106 | 4.048 |
| fish | 12.3659 | 33.1363 | 19.052 | 256.9916 | 38.1686 | 32.8666 | 39.8556 | 52.5336 | 11.697 |
| mhyang | 31.5107 | 51.5048 | 3.6666 | 251.1091 | 9.7712 | 53.9352 | 4.4252 | 43.3887 | 4.741 |
| motocross_2 | 10.2428 | 42.9039 | 31.7723 | 58.2587 | 35.8466 | 55.2349 | 63.9674 | 65.1716 | 14.7628 |
| cup_on_table | 13.8521 | 18.3712 | 1.5787 | 1.9819 | 2.5663 | 14.5659 | 1.8291 | 17.5912 | 1.926 |
| person | 10.4055 | 71.7153 | 68.6291 | 139.2425 | 2.8737 | 14.1563 | 12.9408 | 90.8026 | 5.563 |
| jp1_sequence | 13.6429 | 42.7517 | 4.1736 | 3.2073 | 26.6158 | 14.1563 | 2.3216 | 14.6432 | 6.287 |
| wdesk | 33.4993 | 76.0324 | 89.3332 | 44.1145 | 15.5642 | 45.4719 | 15.3393 | 35.6581 | 14.98 |
| chasing | 12.8025 | 38.6676 | 4.9867 | 16.4593 | 5.8737 | 27.0968 | 6.7806 | 9.3662 | 4.303 |
| can | 13.4378 | 21.9712 | 9.4285 | 30.2203 | 12.7469 | 11.8766 | 12.092 | 32.9759 | 8.095 |
| ped1 | 51.3988 | 66.844 | 59.738 | 364.7516 | 10.3577 | 12.0172 | 64.2581 | 45.1044 | 10.096 |

**Fig. 3**. The average tracking errors. The error is measured using the Euclidian distance of two center points from the ground truth. The best three results are shown in red, blue, and green fonts.

### 3.1. Quantitative Comparison

Fig.3 and Fig.4 show the results of the performance comparison under the two criteria. As we can see, the proposed tracker has overall better performance over others. Even for the sequence where our tracker does not rank first, the performance is still close to the best.

### 3.2. Qualitative Comparison

Fig. 5 shows screenshots of some tracking results.

Low resolution and rotation. In the david2 sequence, the face is in low resolution and undergoes in-plane rotation and out-of-plane rotation. The guided filter enhances the detail of the target. Thus our method can track the target throughout this sequence.

Illumination and scale change. The davidIndoor sequence sequences contain large illumination, pose variation and scale changes. Our decomposition model reconstructs the appearance of the object accurately. Thus alleviate the effects of the illumination change and pose variation.

Rotation and scale changes. The chasing sequence includes scenes with in-plane rotation, out-of-plane rotation and scale changes.

Fast motion and rotation: The can sequence undergoes in-plane rotation and out-of-plane rotation. The appearance of the can changed frequently. The guided filter enhances the property of the can. Our decomposition model reconstructs

| | CT | IVT | L1-APG | L1 | L2-RLS | MIL | MTT-L01 | WMIL | Our1 |
|---|---|---|---|---|---|---|---|---|---|
| basketball | 0.2634 | 0.0152 | 0.2662 | 0.0083 | 0.0938 | 0.2579 | 0.2497 | 0.6234 | 0.941 |
| car | 0.7201 | 0.6768 | 0.2914 | 0.5623 | 0.527 | 0.189 | 0.4385 | 0.0025 | 0.436 |
| car11 | 0.7201 | 0.6768 | 0.9211 | 0.5623 | 0.8295 | 0.3919 | 0.9135 | 0.0025 | 0.863 |
| caviar2 | 0.27 | 0.302 | 0.914 | 0.992 | 0.342 | 0.036 | 0.982 | 0.012 | 0.986 |
| davidIndoor | 0.2359 | 0.2273 | 0.2087 | 0.1991 | 0.2273 | 0.0606 | 0.2857 | 0.2143 | 0.595 |
| dog1 | 0.6022 | 0.2230 | 0.9993 | 0.5659 | 0.99926 | 0.6148 | 0.8652 | 0.2222 | 0.987 |
| faceocc | 0.0011 | 0.0609 | 0.2167 | 0.1309 | 0.1309 | 0.0011 | 0.0158 | 0.0011 | 0.907 |
| faceocc2 | 0.5767 | 0.3877 | 0.4147 | 0.4515 | 0.7067 | 0.5607 | 0.9607 | 0.5546 | 0.954 |
| juice | 0.4703 | 0.349 | 1 | 0.5 | 0.8861 | 0.0074 | 1 | 0.4579 | 0.889 |
| shop | 0.3393 | 0.4036 | 0.9768 | 0.9768 | 0.3643 | 0.00179 | 0.99286 | 0.01250 | 0.984 |
| singer | 0.2906 | 0.3447 | 0.4359 | 0.2393 | 0.0256 | 0.2222 | 0.3476 | 0.2593 | 1 |
| ucsdpeds | 0.5594 | 0.0383 | 1 | 0.0536 | 0.023 | 0.0575 | 0.8352 | 0.0038 | 1 |
| david2 | 0.0019 | 0.0857 | 0.8361 | 0.257 | 0.9646 | 0.00372 | 1 | 0.3259 | 1 |
| fish | 0.9139 | 0.2164 | 0.0735 | 0.0483 | 0.0651 | 0.1639 | 0.042 | 0.0357 | 0.84 |
| mhyang | 0.002 | 0.294 | 0.9913 | 0.2416 | 0.7517 | 0.002 | 1 | 0 | 1 |
| motocross_2 | 0.8696 | 0.3043 | 0.5217 | 0.2609 | 0.3913 | 0.1304 | 0.5652 | 0.0435 | 0.826 |
| cup_on_table | 0.1216 | 0.1863 | 1 | 1 | 1 | 0.1392 | 1 | 0 | 1 |
| person | 0.7434 | 0.1837 | 0.4952 | 0.5111 | 0.9894 | 0.453 | 0.6315 | 0.5322 | 1 |
| jp1_sequence | 0.8569 | 0.0872 | 0.7763 | 1 | 0.699 | 0.453 | 1 | 0.9441 | 1 |
| wdesk | 0.5021 | 0.1171 | 0.0621 | 0.3738 | 0.5501 | 0.0033 | 0.6756 | 0.6601 | 0.985 |
| chasing | 0.1783 | 0.0667 | 0.64 | 0.7383 | 0.6717 | 0.8983 | 0.005 | 0.705 | 0.97 |
| can | 0.6173 | 0.2588 | 0.7305 | 0.4286 | 0.6092 | 0.7008 | 0.6846 | 0.1078 | 0.821 |
| ped1 | 0.5897 | 0.4316 | 0.2692 | 0.0085 | 0.5897 | 0.9188 | 0.9615 | 0.5385 | 0.987 |

**Fig. 4**. Average overlap rate. The best three results are shown in red, blue, and green fonts.

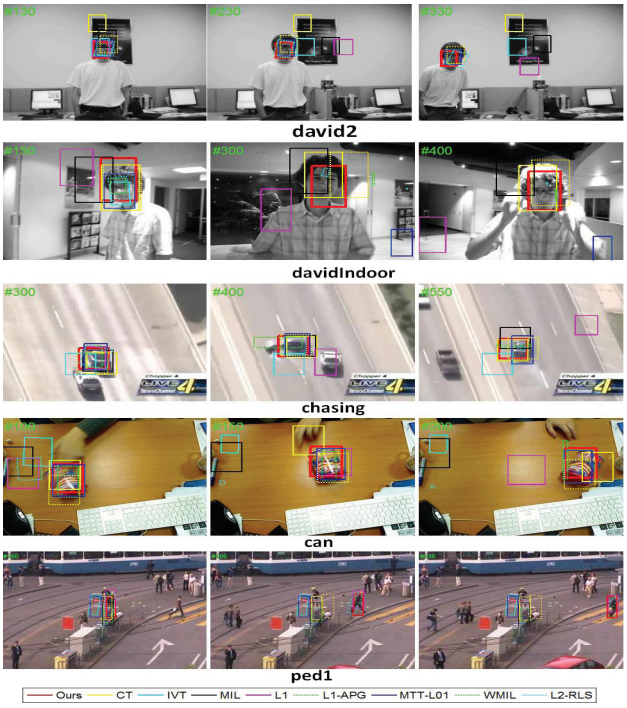


**Fig. 5**. Shows screenshots of some tracking results.

the varied appearance of the target. Therefore, our tracker achieves the best performance.

Occlusion and low resolution: In the ped1 sequence, the target is severely occluded. The guided filter preserves the property of targets in the low resolution. The context information helps the tracker to locate the target when the occlusion happened.

Overall, our proposed tracker performs well against state-of-the-art methods on these challenging sequences, which can be attributed to three main reasons. First, target objects are represented by the target and context information. The proposed representation includes the context is able to distinguish the target from the background. And the context information prevents the drift happens. Second, the guided filter act on subspace which enable our tracker less sensitive to large pose variations and background clutter (daved2, can and cup on table) and blurring caused by fast motion (motocross2). Third, the multi-task sparse learning which considers the candidates jointly is able to handle the challenges of significant illumination (chasing, davidIndoor, fish and david2) and severe occlusion (ped1, faceocc and faceocc2).

## 4. CONCLUSION

In this paper, we propose an innovative algorithm for visual object tracking. Our method considers the context information. Owing to the guided filter, the candidates can be represented compactly be the dictionary template. We further model the appearance representation by multi-task sparse learning to generate a robust tracker. Extensive experimental results show that the proposed algorithm performs favorably against the state-of-the-art methods in terms of accuracy and robustness.

## 5. REFERENCES

[1] S. K. Divvala, D. Hoiem, J. H. Hays, A. A. Efros, and M. Hebert, An empirical study of context in object detection, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Miami, FL, USA, 2009, pp. 1271C1278.

[2] D. Munoz, J. A. Bagnell, N. Vandapel, and M. Hebert, Contextual classification with functional max-margin Markov network, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Miami, FL, USA, 2009, pp. 975C982.

[3] M. Ozuysal, P. Fua, and V. Lepetit, Fast keypoint recognition in ten lines of code, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Minneapolis, MN, USA, 2007, pp. 1C8.

[4] Chakravorty, Tanushri, Guillaume-Alexandre Bilodeau, and Eric Granger. Contextual Object Tracker with Structure Encoding, ICIP, 2015

[5] Zhang T, Ghanem B, Liu S, et al. Robust Visual Tracking via Exclusive Context Modeling[J]. 2015.

[6] Zhang K, Zhang L, Yang M H, et al. Fast tracking via spatio-temporal context learning[J]. arXiv preprint arXiv:1311.1939, 2013.

[7] Dandan Du, Huchuan Lu, Lihe Zhang, Fu Li, Visual Tracking Via Guided Filter, International Conference on Image Processing, 2015

[8] Zhang, T., Ghanem, B., Liu, S., Ahuja, N., Robust visual tracking via multi-task sparse learning. In IEEE conference on computer vision and pattern recognition (pp. 1C8), (2012)

[9] Ali Jalali, Pradeep Ravikumar, Sujay Sanghavi and Chao Ruan, A Dirty Model for Multitsak Learning, NIPS, 2010

[10] Wang, Yong, Shiqiang Hu, and Shandong Wu, Visual tracking based on group sparsity learning, Machine Vision and Applications, pp: 1-13, 2014.

[11] Mei, X., Ling, H., Robust visual tracking and vehicle classification via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(11), 2259-2272, (2011)

[12] K. He, J. Sun, and X. Tang. Guided image filtering. TPAMI, 35(6):1397-1409, 2013.

[13] Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J, Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn., 3(1):1-122, (2011)

[14] X. Chen, W. Pan, J. Kwok, and J. Carbonell, Accelerated gradient method for multi-task sparse learning problem. In IEEE international conference on data mining (pp. 746-751), (2009)

[15] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, The PASCAL Visual Object Classes Challenge 2010 (VOC2010) Results, (2010)

[16] Ross, D., Lim, J., Lin, R.S., Yang, M.H., Incremental learning for robust visual tracking. International Journal of Computer Vision, 77(1), 125C141, (2008)

[17] C. Bao, Y. Wu, H. Ling, and H. Ji, Real Time Robust L1 Tracker Using Accelerated Proximal Gradient Approach, IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Rhode Island, (2012)

[18] Boris Babenko, Ming-Hsuan Yang, and Serge Belongie, Robust Object Tracking with Online Multiple Instance Learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(8), 1619-1632, (2011)

[19] Kaihua Zhang, Lei Zhang, and Ming-Hsuan Yang, Real-Time Compressive Tracking. Proceedings of European Conference on Computer Vision, vol. 3, pp. 864-877, Florence, Italy, October, (2012)

[20] Ziyang Xiao, Huchuan Lu, Dong Wang: L2-RLS-Based Object Tracking. IEEE Trans. Circuits Syst. Video Techn. 24(8): 1301-1309 (2014)

[21] Zhang K, Song H. Real-time visual tracking via online weighted multiple instance learning [J]. Pattern Recognition, 2013, 46(1): 397-411.