A Simplified Low Rank and Sparse Model for Visual Tracking

Mi Wang, Huaxin Xiao, Yu Liu, Wei Xu and Maojun Zhang
Institute of Information Systems and Management, National University of Defense,
No. 109 Deya Road Kaifu District Changsha City Hunan Prov., 410073, Changsha, China
{coolsa007, huaxin_xiao}@163.com, jasonyuliu@nudt.edu.cn, {545406612, 1023133126}@qq.com

Keywords: Visual Tracking, Sparse and Low-Rank Representation.

Abstract: Object tracking is the process of determining the states of a target in consecutive video frames based on properties of motion and appearance consistency. Numerous tracking methods using low-rank and sparse constraints perform well in visual tracking. However, these methods cannot reasonably balance the two characteristics. Sparsity always pursues a sparse enough solution that ignores the low-rank structure and vice versa. Therefore, this paper replaces the low-rank and sparse constraints with $l_{2,1}$ norm. A simplified low-rank and sparse model for visual tracking (LRSVT), which is built upon the particle filter framework, is proposed in this paper. The proposed method first prunes particles which are different with the object and selects candidate particles for efficiency. A dictionary is then constructed to represent the candidate particles. The proposed LRSVT algorithm is evaluated against three related tracking methods on a set of seven challenging image sequences. Experimental results show that the LRSVT algorithm favorably performs against state-of-the-art tracking methods with regard to accuracy and execution time.

1 INTRODUCTION

Visual tracking finds a region in the current image that matches the given object. It is a well-known problem in computer vision with numerous applications including surveillance, driver assistance, robotics, human-computer interaction, and motion analysis (Zhang T et al. 2014). Despite demonstrated success, it remains challenging to design a robust visual tracking algorithm due to factors such as occlusion, background clutter, varying viewpoints, and illumination and scale changes (Wang L et al. 2015).

Recently, sparse and low-rank representation has cause for concern in many aspects (R. Xia et al. 2014, Zhang C et al. 2015). These tracking methods express a target by a sparse linear combination of the templates in a dictionary (Zhang T et al. 2014). These algorithms based on l_1 minimization have been demonstrated time-consuming. Then they set up low-rank representation and sparse representation to solve the problem. However, they can not balance the two characteristics in good reason. Sparse always pursue a sparse enough solution, which ignoring the low-rank structure. At the same time, $l_{2,1}$ norm has been proved effective at represent both low-rank and sparse in some paper (Zhao M et al. 2014). Besides, the $l_{2,1}$ norm avoid the time-consuming process of nuclear norm.

This paper, we use norm which can combine the low-rank and sparse characteristic to learn robust linear representations for efficient and effective object tracking. The proposed visual tracking algorithm is developed based on the particle filter. We can see the process in Fig. 1.

Fig. 1 shows the flowchart of the enforcement of the proposed algorithm by pruning particles. First, the target is selected from the first frame. Second, all particles are sampled based on the previous object. Third, the particles are pruned using the reconstruction error to prune particles. Finally, the object is selected using our LRSVT algorithm in the next frame, which enforces sparsely low-rank properties.

Wang, M., Xiao, H., Liu, Y., Xu, W. and Zhang, M.
A Simplified Low Rank and Sparse Model for Visual Tracking.
DOI: 10.5220/0006117003010308
In *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2017)*, pages 301-308
ISBN: 978-989-758-222-6

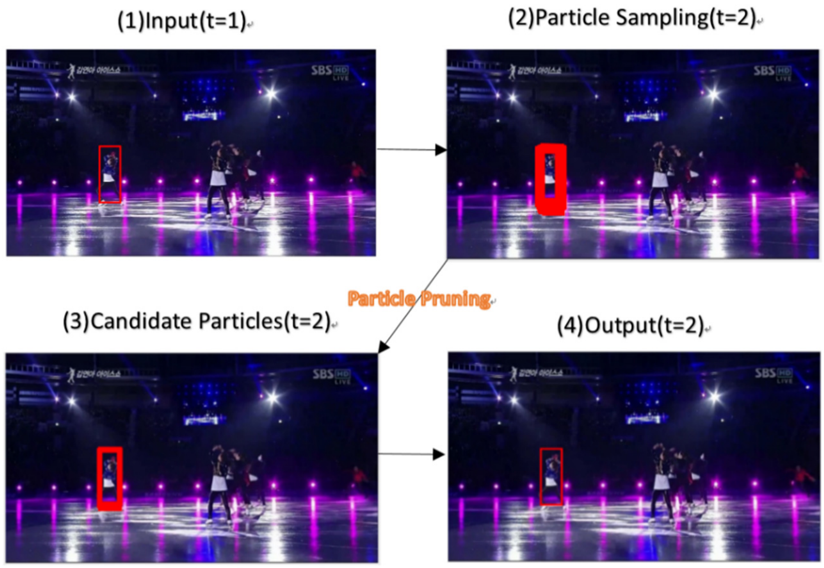

Figure 1: Enforcing the sparsity, low-rank properties in the proposed LRSVT algorithm. (1) The frame at time t(t=1). (2) All particles sampled based on previous object. Here the number of particle is = 400. (3) Particles are pruned using the reconstruction error e_0 . 25 candidate particles are obtained after pruning. (4) The frame at time t(t=2), object is selected using our LRSVT algorithm in the next frame.

Object tracking is formulated as a sparse and low-rank representation problem from a new perspective, which is carried out by exploiting the relationship between the observations of the particle samples and jointly representing them using a dictionary of templates with an online update. The resulting sparsely low-rank representation of candidate particles facilitates robust performance for visual tracking. The relationship of these algorithms and the importance of each property for visual tracking are shown.

2 RELATED WORKS

The recent years have witnessed significant progress in tracking with sparse and low-rank representation. Most recently, an algorithm that jointly learns the sparse and low-rank representations of all particles (Zhang, K. et al. 2012; Zhang, T. et al. 2012–2014) is proposed for object tracking. Solutions to low-rank matrix minimization and completion problems have also achieved considerable progress. Zhou X et al (Zhou X et al. 2012) demonstrated that the image sequence of a cardiac cycle can be well approximated with a low-rank matrix. Zhang C (Zhang C et al. 2014) learned the observation model by extracting low-rank features. Yehui Yang et al (W Hu et al. 2016) developed a comprehensive study of the $l_{2,1}$ norm to tolerate the sudden changes between two adjacent frames that exploits the low-rank structure among consecutive target observations.

3 LOW RANK SPARSE VISUAL TRACKING

In this section, we present the proposed tracking algorithm based on low-rank sparse representations of particle samples.

3.1 Consistent Low-rank Sparse Representation

In this work, particles are sampled from previous object locations to predict the state s_t of the target at time t, from which the region of interest y_t is cropped in the current image and normalized to the template size. The state transition function $p(s_t | s_{t-1})$ is modeled by an affine motion model with a diagonal Gaussian distribution. The observation model $p(y_t | s_t)$ reflects a similarity between an observed image region y_t corresponding to a particle s_t and the templates of the current dictionary. In this paper, $p(y_t | s_t)$ is computed as a function of the difference between the consistent low-rank sparse representation of the target based on object templates and its representation based on background templates. The particle that maximizes this function is selected as the tracked target at each time instance. At time t, n_0 sampled particles and corresponding vectorized gray-scale image observations form a matrix $X_0 = [x_1, x_2, ..., x_{n_0}]$, wherein the observation with regard to the i-th particle is denoted as $x_i \in R^d$. Each observation is represented as a linear combination of templates from a dictionary $D_t = [d_1, d_2, ..., d_m]$, such that $X_0 = D_t Z_t$. The columns of $Z_t = [z_1, z_2, ..., z_{n_0}]$ denote the representations of particle observations with regard to D_t. The dictionary columns contain templates used to represent each particle, including image observations of the tracked object and the background. Misalignment between dictionary templates and particle observations may lead to tracking drifts because representation is constructed on the pixel level. The dictionary D_t can be constructed from an over-complete set using transformed templates of the target and background classes to alleviate this problem. This dictionary is also updated progressively. Temporal consistency is exploited to prune particles for efficient and effective tracking. A particle is considered temporally inconsistent if its observation is not linearly well represented by the dictionary D_t and the representation of the tracked target in the previous frame, which is denoted as z_0. More specifically, the particle is pruned in the current frame if the l_2 reconstruction error $\|x_i - D_t z_0\|_2$ is sufficiently large, thereby leaving a number of x_i; therefore, the number is set as n. In this work, temporal consistency is exploited as the appearances of the tracked object.Consequently, this process effectively reduces the number of particles to be represented from n_0 to n, where $n_0 \gg n$ in most cases. Next, the ones left after pruning are denoted as candidate particles, in which their corresponding observations are $X \in R^{d \times n}$ and their representations are $Z \in R^{m \times n}$.

The representation of each candidate particle is based on the following observations. (1) After pruning, the candidate particle observations can be modeled by a low-rank subspace (i.e., X is low-rank); therefore, Z (i.e., their representations with regard to D_t) is expected to be low-ranked. (2) The observation x_i of a good candidate particle can be modeled by a small number of nonzero coefficients in its corresponding representation z_i. (3) The aim of object tracking is to search for patches (with regard to particles) that have a representation similar to previous tracking results. Therefore, a "good" representation should be consistent over time. In the work of CLRST [4], the tracking problem is formulated by min Z, E

$$\min_{Z,E} \lambda_1 \|Z\|_* + \lambda_2 \|Z\|_{1,1} + \lambda_3 \|Z - Z_0\|_{2,1} + \lambda_4 \|E\|_{1,1} \quad (1)$$
$$s.j. \quad X = DZ + E$$

where

$$\|Z\|_{p,q} = \left(\sum_j \left(\sum_i \left| [Z]_{ij} \right|^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}} \quad (2)$$

$\lambda_1 \|Z\|_* + \lambda_2 \|Z\|_{1,1}$ as $\lambda_1 \|Z\|_{2,1}$ is replaced in this paper.

The $l_{2,1}$ norm encourages the columns of Z to be zero, which assumes that the corruptions are "sample-specific" (i.e., several data vectors are corrupted and the others are clean) (Zhang X et al. 2012) to ensure that Z has a low-rank and sparse property.

$$\min_{Z,E} \lambda_1 \|Z\|_{2,1} + \lambda_3 \|Z - Z_0\|_{2,1} + \lambda_4 \|E\|_{1,1} \quad (3)$$
$$s.j. \quad X = DZ + E$$

E is the error which is attributed to noise as well as occlusion.

We then lead in two equality constraints, and the equation and constraint becomes

$$\min_{Z,E} \lambda_1 \|Z_1\|_{2,1} + \lambda_3 \|Z_2\|_{2,1} + \lambda_4 \|E\|_{1,1}$$
$$s.j. \begin{cases} X = DZ_3 + E \\ Z_3 = Z_1 \\ Z_3 = Z_2 + Z_0 \end{cases} \quad (4)$$

In this formulation, λ_i, i = 1, 3, 4 are weights that quantify the trade-off between the different terms discussed below. In addition, $[Z]_{ij}$ denotes the entry at the i-th row and j-th column of Z. The representation of the previous tracking result is denoted with regard to D_t as z_0. The matrix $Z_0 = 1 z_0^{\mathrm{T}}$ is a rank one matrix, where each column is z_0.

3.1.1 Low-Rank and Sparse: $\|Z\|_{2,1}$

In CLRST formulation, $\|Z\|_*$ is used to minimize the matrix rank of representations of all candidate particles together. Their sparse representation scheme is $\|Z\|_{1,1}$, which has been shown to be robust to occlusion or noise in visual tracking. $\|Z\|_{2,1}$ is considered to replace $\lambda_1\|Z\|_* + \lambda_2\|Z\|_{1,1}$ which is the sparse congruency constraint on matrix Z. This constraint only allows a few rows of Z to become nonzero, thereby deleting the ambiguous bases and maintaining principal bases. Therefore, the samples belonging to the same class are more likely to choose the same atom in their representation and share the same sparse pattern in their SR coefficient vectors. Thus, Z is sparse and low-rank. By contrast, the sparse congruency constraint considers the global structure of Z and eliminates rows of elements that have a slight contribution to the representation of the dataset and do not affect the low-rank structure of Z. Thus, the contribution time is greatly reduced (Zhao M et al. 2014).

3.1.2 Temporal $\|Z-Z_0\|_{2,1}$ and Reconstruction Error $\|E\|_{1,1}$

Temporal representation allows only a small number of particles to have representations different from the previous tracking results. The values and support of the columns in E are informative because these values indicate the presence of occlusion (substantial values but sparse support) and determines whether a candidate particle is sampled from the background (substantial values with non-sparse support) (Zhao M et al. 2014).

3.2 Solving

3.2.1 Solving Equation

$$\begin{aligned}
&L\left(Z_{1,2,3}, E, Y_{1,2,3}, u_{1,2,3}\right) \\
&= \lambda_1\|Z_1\|_{1,2} + \lambda_3\|Z_2\|_{2,1} + \lambda_4\|E\|_{1,1} \\
&+tr\left[Y_1^T\left(X - DZ_3 - E\right)\right] + \frac{u_1}{2}\|X - DZ_3 - E\|_F^2 \\
&+tr\left[Y_2^T\left(Z_3 - Z_1\right)\right] + \frac{u_2}{2}\|Z_3 - Z_1\|_F^2 \\
&+tr\left[Y_3^T\left(Z_3 - Z_2 - Z_0\right)\right] + \frac{u_3}{2}\|Z_3 - Z_2 - Z_0\|_F^2
\end{aligned} \tag{5}$$

3.2.2 Solving Z_1, Z_2, E, Z_3 in Turn

$$\begin{aligned}
Z_1^* &= \arg\min \frac{\lambda_1}{u_2}\|Z_1\|_{1,2} + \frac{1}{2}\left\|Z_1 - Z_3 - \frac{1}{u_2}Y_2\right\|_F^2 \\
&= \mathcal{L}'_{\frac{\lambda_1}{u_2}}\left(Z_3 + \frac{1}{u_2}Y_2\right)
\end{aligned} \tag{6}$$

$$\begin{aligned}
Z_2^* &= \arg\min \frac{\lambda_3}{u_3}\|Z_2\|_{2,1} + \frac{1}{2}\left\|Z_2 - Z_3 + Z_0 - \frac{1}{u_3}Y_3\right\|_F^2 \\
&= \mathcal{L}_{\frac{\lambda_3}{u_3}}\left(Z_3 - Z_0 + \frac{1}{u_3}Y_3\right)
\end{aligned} \tag{7}$$

$$\begin{aligned}
E^* &= \arg\min \frac{\lambda_4}{u_1}\|E\|_{1,1} + \frac{1}{2}\left\|E - X + DZ_3 - \frac{1}{u_1}Y_1\right\|_F^2 \\
&= \mathcal{S}_{\frac{\lambda_4}{u_1}}\left(X - DZ_3 + \frac{1}{u_1}Y_1\right)
\end{aligned} \tag{8}$$

And

$$\begin{aligned}
Z_3^* &= \arg\min tr\left[Y_1^T\left(X - DZ_3 - E\right)\right] \\
&+ \frac{u_1}{2}\|X - DZ_3 - E\|_F^2 \\
&+tr\left[Y_2^T\left(Z_3 - Z_1\right)\right] + \frac{u_2}{2}\|Z_3 - Z_1\|_F^2 \\
&+tr\left[Y_3^T\left(Z_3 - Z_2 - Z_0\right)\right] + \frac{u_3}{2}\|Z_3 - Z_2 - Z_0\|_F^2 \\
&= G_1\left[D^T\left(X - E\right) + G_2 + G_3\right]
\end{aligned} \tag{9}$$

Where

$$G_1 = \left(D^T D + \frac{u_2}{u_1}I + \frac{u_3}{u_1}I\right)^{-1} \tag{10}$$

$$G_2 = \frac{u_2}{u_1}Z_1 + \frac{u_3}{u_1}\left(Z_2 + Z_0\right) \tag{11}$$

and

$$G_3 = \frac{1}{u_1}\left(D^T Y_1 - Y_2 - Y_3\right) \tag{12}$$

3.2.3 Update $Y_{1,2,3}, u_{1,2,3}$

$$\begin{cases}
Y_1 = Y_1 + u_1\left(X - DZ_3 - E\right) \\
Y_2 = Y_2 + u_2\left(Z_3 - Z_1\right) \\
Y_3 = Y_3 + u_3\left(Z_3 - Z_2 - Z_0\right) \\
u_1 = \rho u_1; u_2 = \rho u_2; u_3 = \rho u_3;
\end{cases} \tag{13}$$

3.3 Adaptive Dictionary

The dictionary D_t is initialized by sampling image patches around the initial target position. The dictionary is updated in successive frames to model the change in appearance of the target object and to ensure accuracy in the tracking. D_t is augmented with representative templates of the background to alleviate the problem of tracking drift, such that $D_t = [D_O\ D_B]$, where D_O and D_B represent the target object and background templates, respectively. Thus, the representation z_k of a particle comprised an object representation z_k^O and a background representation z_k^B. The tracking result y_t at instance t is the particle x_i, such that

$$\mathrm{i} = \arg\max_{k=1,\dots,n} \left(\left\| \mathrm{z}_\mathrm{k}^O \right\|_1 - \left\| \mathrm{z}_\mathrm{k}^B \right\|_1 \right) \tag{14}$$

which encourages good modeling of the tracking result using object templates and not using background templates. Discriminative information was also employed to design a systematic procedure for updating D_t.

4 EXPERIMENT

In this section, the experimental results on the evaluation of the proposed tracking algorithm against several state-of-the-art methods were evaluated.

4.1 Datasets

Twenty-five challenging videos with ground truth object locations, including *basketball, football, singer1, singer2, singer1(low frame rate), skating1,* and *skating2* were used for analysis. These videos contain complex scenes with challenging factors (e.g., cluttered background, moving camera, fast movement, large variation in pose and scale, occlusion, shape deformation, and distortion).

4.2 Evaluated Algorithms

The proposed tracking methods (SLRVT) are compared with three state-of-the-art visual trackers, including FCT (Zhang K et al. 2014), l_1 (Zhao M et al. 2014), and CLRST (Mei X et al. 2011). Publicly available sources or binary codes provided by the authors are used for fair comparisons. The same initialization and parameter settings in all experiments are also used.

4.3 Evaluation Criteria

Two metrics are used to evaluate tracking performance. The first metric is the center location error, which is the Euclidean distance between the central location of a tracked target and the manually labeled ground truth. The second metric is an overlap ratio based on the PASCAL challenge object detection score (Everingham B M. et al. 2010). Given tracked bounding box ROI_T and the ground truth bounding box ROI_{GT}, the overlap score can be computed as

$$score = \frac{area(ROI_T \cap ROI_{GT})}{area(ROI_T \cup ROI_{GT})} \tag{15}$$

The average overlap score across all frames of each image sequence is computed to rank the tracking performance.

4.4 Implementation Details

All experiments are carried out in MATLAB on a 3.2 GHz Intel Corei5–4460 Duo machine with 4 GB RAM. Template size d, which is manually initialized in the first frame, is set to half the size of the target object. The affine transformation, where the state transitional probability $p(y_t \mid s_{t\text{-}1})$ is modeled by a zero-mean Gaussian distribution and a diagonal covariance matrix σ_0 with values (0.03, 0.0005, 0.0005,0.03, 1, 1): $p(y_t \mid s_{t\text{-}1}) \sim N(0, \sigma_0)$, is used. The definition of $p(y_t \mid s_t)$ is $p(y_t \mid s_t) = \vartriangle \mathrm{z}_\mathrm{i} (i = 1, 2, \dots, n)$. The representation threshold is set to 0.5. Parameter σ is set to 1.0 in the CLRST method to prune particles. The number of particles n_0 is set to 400 and total number of templates m is set to 25.

5 TEST RESULTS

5.1 Parameter Analysis

Several parameters play important roles in the proposed tracking algorithm. In this section, determining the values and effects of these parameters on tracking performance is shown.

Effect of λ :

The objective function has three parameters, namely,

Table 1: a. The distance value with the change of λ_1. b. The distance value with the change of λ_3.

a. The distance on the value of λ_1

λ_1:	0.0001	0.5	0.9	1	1.1	2	5	10
Distance	50.1	54	50.6	33.7	54.3	53.2	55.8	58.4

b. The distance on the value of λ_3

λ_3:	0.0001	0.5	0.9	1	1.1	2	5	10
Distance	55.5	54	51.5	41.7	54.6	44.8	52	51.1

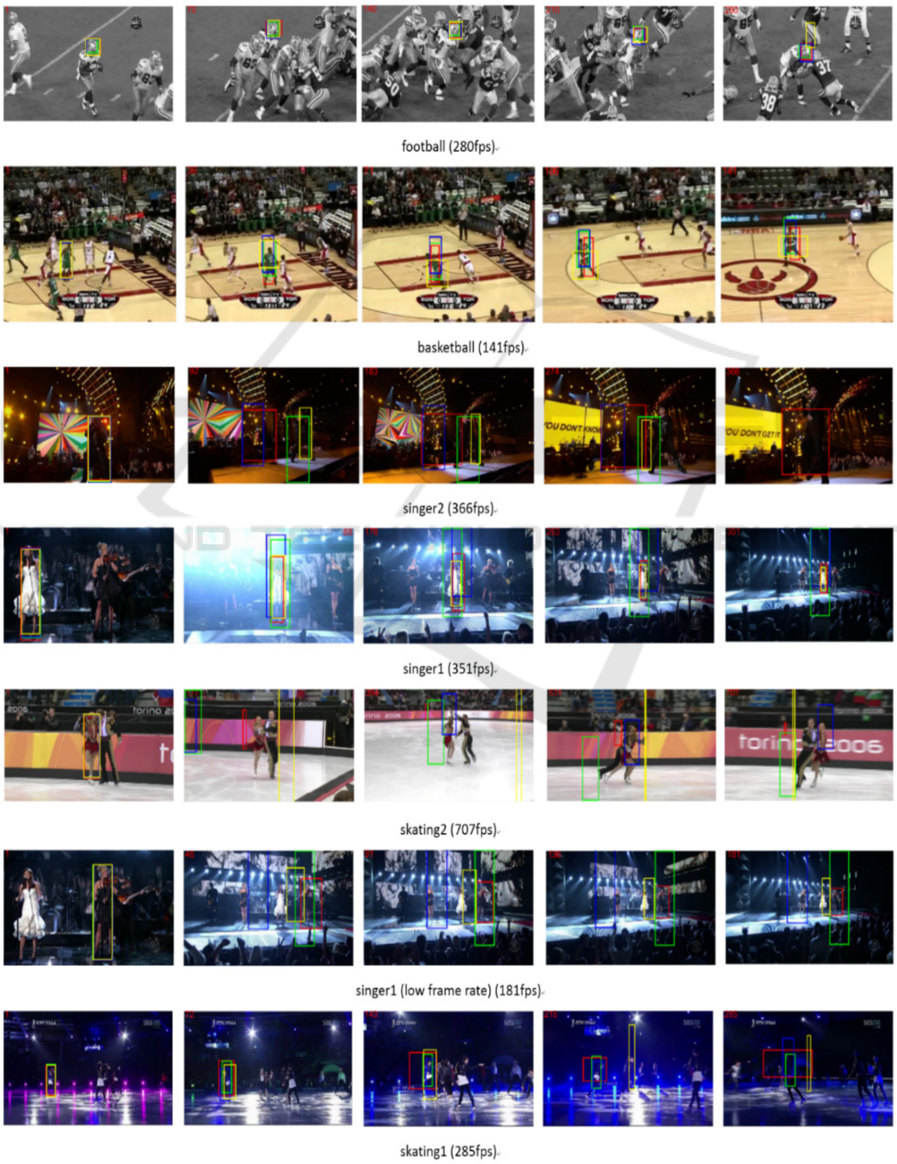

Figure 2: Tracking result on 7 image sequences. LRSVT、FCT (Zhang K et al. 2014)、CLRST (Zhao M et al. 2014)、l_1 (Mei X et al. 2011)are respectively displayed in *red*、*green*、*blue* and *yellow.*

Table 2: Overlap、distance and time result on 7 image sequences with LRSVT、CLRST、FCT and l_1 methods.

a. Overlap

	LRSVT	CLRST	l_1	FCT
football(280fps)	0.262	0.312	0.582	0.283
basketball(141fps)	0.381	0.506	0.661	0.501
singer2(366fps)	0.436	0.588	0.408	0.328
singer1(351fps)	0.436	0.588	0.818	0.047
skating2(707fps)	0.284	0.17	0.101	0.255
singer1(low frame rate) (181fps)	0.592	0.095	0.344	0.255
skating1(285fps)	0.511	0.642	0.411	0.669

b. Distance

	LRSVT	CLRST	l_1	FCT
football(280fps)	5.2	4.2	11.4	6
basketball(141fps)	9	23.3	23	16.5
singer2(366fps)	130	133.7	35.1	49.7
singer1(351fps)	130	133.7	3.5	22.4
skating2(707fps)	98.4	64.2	184.8	127.6
singer1(low frame rate)(181fps)	4.7	162.7	37	127.6
skating1(285fps)	24.4	8.3	35.9	18.3

c. Time

	LRSVT	CLRST	l_1	FCT
football(280fps)	0.431	1.597	0.00027	0.0178
basketball(141fps)	0.563	2.354	0.000269	0.0195
singer2(366fps)	0.59	2.555	0.000089	0.0174
singer1(351fps)	0.59	2.555	0.000136	0.0186
skating2(707fps)	0.821	1.608	0.000085	0.019
singer1(low frame rate) (181fps)	0.581	2.65	0.000194	0.019
skating1(285fps)	0.586	2.674	0.000216	0.018

λ_1, λ_3, and λ_4 (1). Because λ_1 and λ_3 are related to the coefficients Z and λ_4 is related to E, $\lambda_4 = 1$ was fixed and other parameter values were changed. All $\lambda_i\,(i=1,3)$ are parameterized by a discrete set Λ for sensitivity analysis, in which Λ = {0.0001, 0.5, 0.9,1, 1.1, 2, 5, 10.0}. The different combinations of these values were analyzed on video with 100 frames. The average distance score from all frames was computed for each combination. The corresponding results were obtained for different λ_1, as shown in Table 1.a. Table 1 shows the sensitivity analysis of $\lambda_i\,(i=1,3)$. From on these results, we can set $\lambda_1 = 1$, $\lambda_3 = 1$, and $\lambda_4 = 1$ for the objective function (1).

5.2 Qualitative Comparison

Fig. 2 and Table 2 show the tracking results of four trackers on seven sequences. Three norms are included: overlap, distance, and time.

Singer1(low frame rate) has better tracking performance based on the visual effect of the views of *football, basketball,* and *singer1*. The proposed method performed well in terms of position and size of the target. The *singer2* sequence contains significant illumination, scale, and viewpoint changes. *skating2* contains Abrupt Motion, Illumination Change, and Occlusion. Therefore, most trackers drift away from the target object in these two sequences. In the *Singer2* sequence, only the result of the LRSVT method falls on the screen. In *Skating1* sequence, length and width did not fully track the target in terms of the basic location of the tracking target.

LRSVT performed well at overlap in *singer1(low frame rate)* and at the distance in *basketball* than any of the other methods. Among all sequences, the time consumed from fastest to slowest is in the order of l_1 , FCT, LRSVT, and CLRST.

6 CONCLUSION

This paper conducted based on the CLRST method. $l_{2,1}$ norm was used to represent low-rank and sparse, which differs from CLRST. The performance of the tracking algorithms against three competing state-of-the-art methods on seven challenging image sequences was analyzed extensively. The proposed method significantly reduced computation time than CLRST. The result maintained more than twice the speed of operation with the same overlap and distance. The results are in line with expectations.

ACKNOWLEDGEMENT

This research was partially sponsored by National Natural Science Foundation (NSFC) of China under project No.61403403 and No.61402491.

REFERENCES

Zhang, T., Liu, S., Ahuja, N., Yang, M. H., & Ghanem, B. (2014). Robust visual tracking via consistent low-rank sparse learning. International Journal of Computer Vision, 111(2), 171-190.

Shen, Z., Toh, K. C., & Yun, S. (2011). An accelerated proximal gradient algorithm for frame-based image restoration via the balanced approach. Siam Journal on Imaging Sciences, 4(2), 573-596.

Zhang, K., Zhang, L., & Yang, M. H. (2012). Real-Time Compressive Tracking. European Conference on Computer Vision (Vol.7574, pp.864-877). Springer-Verlag.

Zhang, T., Ghanem, B., & Ahuja, N. (2012). Robust multi-object tracking via cross-domain contextual information for sports video analysis. , 22(10), 985-988.

Zhang, T., Ghanem, B., Xu, C., & Ahuja, N. (2013). Object tracking by occlusion detection via structured sparse learning. , 71(4), 1033-1040.

Zhang, T., Ghanem, B., Liu, S., & Ahuja, N. (2012). Robust visual tracking via multi-task sparse learning. , 157(10), 2042-2049.

Zhang, T., Ghanem, B., Liu, S., & Ahuja, N. (2013). Robust visual tracking via structured multi-task sparse learning. International Journal of Computer Vision, 101(2), 367-383.

Zhang, T., Ghanem, B., Liu, S., Xu, C., & Ahuja, N. (2013). Low-Rank Sparse Coding for Image Classification. IEEE International Conference on Computer Vision (pp.281-288).

Zhang, T., Ghanem, B., Xu, C., & Ahuja, N. (2013). Object tracking by occlusion detection via structured sparse learning. , 71(4), 1033-1040.

Zhang, X., Ma, Y., Lin, Z., Gao, H., Zhuang, L., & Yu, N. (2012). Non-negative low rank and sparse graph for semi-supervised learning. IEEE Conference on Computer Vision & Pattern Recognition (Vol.157, pp.2328 - 2335).

Zhang, T., Ghanem, B., Liu, S., & Ahuja, N. (2012). Low-Rank Sparse Learning for Robust Visual Tracking. Computer Vision – ECCV 2012. Springer Berlin Heidelberg.

Zhao, M., Jiao, L., Feng, J., & Liu, T. (2014). A simplified low rank and sparse graph for semi-supervised learning ☆. Neurocomputing, 140(Supplement 1), 84-96.

Zhang, K., Zhang, L., & Yang, M. H. (2014). Fast compressive tracking. IEEE Transactions on Pattern Analysis & Machine Intelligence, 36(10), 2002-15.

Everingham, M., Zisserman, A., Williams, C. K. I., Van Gool, L., Allan, M., & Bishop, C. M., et al. (2006). The 2005 pascal visual object classes challenge. Lecture Notes in Computer Science, 111(1), 117-176.

Wang, L., Ouyang, W., Wang, X., & Lu, H. (2015). Visual Tracking with Fully Convolutional Networks. IEEE International Conference on Computer Vision (pp.3119-3127). IEEE.

Zhang, C., Fu, H., Liu, S., Liu, G., & Cao, X. (2015). Low-Rank Tensor Constrained Multiview Subspace Clustering. IEEE International Conference on Computer Vision. IEEE.

Xia, R., Pan, Y., Du, L., & Yin, J. (2014). Robust multi-view spectral clustering via low-rank and sparse decomposition. Twenty-Eighth AAAI Conference on Artificial Intelligence. AAAI Press

Mei, X., Ling, H., Wu, Y., Blasch, E., & Bai, L. (2011). Minimum error bounded efficient l1 tracker with occlusion detection (preprint). Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1257-1264.

Zhou, X., Yang, C., & Yu, W. (2012). Moving object detection by detecting contiguous outliers in the low-rank representation. *IEEE Transactions on Software Engineering, 35*(3), 597-610.

Hu, W., Yang, Y., Zhang, W., & Xie, Y. (2016). Moving object detection using tensor based low-rank and saliently fused-sparse decomposition. *, PP*(99), 1-1.