ISSN: 2688-8653

Realtime Fusion Tracking Infrared-visible Target Using L1-APG

Haifei Ma, Lei Yang, Yong Fan

Guangdong University of Science and Technology 523083, China

Abstract

Aimed at the problem of poor stability within the tracking system of single sensor, this paper addresss a realtime fusion tracking approach based on L1-APG for Infrared-visible target. Firstly, the target models of infrared and visible are built respectively by using sparse representation method, and the optimization problem is constructed by minimizing their joint reconstruction error. Secondly, the optimization problem is solved by employing L1-APG algorithm. Finally, the computational complexity of the algorithm is further reduced by using the minimum error boundary constraint. In addition, an online update method for Infrared-visible target model is designed. Experimental evaluations on several Infrared-visible image sequences demonstrate that the proposed fusion tracker performs well in handling object occlusion, rotation, intersection and night illumination.

Keywords

Fusion Tracking; Infrared-visible Target; Sparse Representation; L1-Accelerated Proximal Gradient (L1-APG).

1. Introduction

Target tracking is a major foundation in the computer vision .such as visual surveillance, human machine interaction, vehicle navigation, video scene intelligent analysis, the majority of tracking methods have been reported [1]. these method are divided into two sections.one is single-sensor ,another is mul-sensor,At present,First of all,Single sensor have an vulnerable impact on the natural environment, then it is too weak in the limited information and efficency to track for the long time, even there is the defect that the narrower spatial and temporal coverage are gained and lower reliability is measured.multi-sensor, nevertheless,in the respect has an inherent advantage, the reason why is widespreadly applied for the target tracking. in which the most typical is a fusion of infrared and visible trace. Generally There are two categories generalized single senior and multiple senior,Single senior is commonly involved random method[2-3], Mean shift determination method[4-5], classification methods [6-8], and subspace methods[9-10]. Since the infrared sensor can form an image by detecting the difference among the target reflecting heat, under poor illuminating conditions or in cases in which the target and the background are the same color, but it can't sense the color information and texture features. The visible sensor, exactly, makes up for these defects.which is obtained higher revolution and achieved spatial scence details from target object.When multiple target is intersecting,we can recognize the target depending on the difference of color information and feature texture. Therefore,the majority of people,joining the two data fusion tracking Infrared-visible target ,can obtain better tracking performance than the single senor. Sun et al. [11] presented a beneficial method to fusion tracking in color and infrared images using joint sparse representation.Cheng et al.[12] introduced an approach of visible/ infrared dual-channel target tracking based on weighted mean-shift,which remained the tracking perform by calculating with Bhattacharyya coefficient corresponding the weight in each channel. Zhao et al.[13]employed the multi-feature fusion to improve the performance

ISSN: 2688-8653

infrared and visible target tracking, Yun et al. [14] proposed a compressive tracking based on time-space Kalman fusion model,which forms a fresh fusion model.

Aimed at the above problems within the tracking system of infrared and visible target joint sparse representation. The paper, therefore in the particle frame,presents a realtime fusion tracking approach using L1-Accelerated Proximal Gradient(shortening L1-APG) for Infrared-visible target. That is to say, the appearance model of target can be described by joint sparse representation and applied 1-norm to calculate sparse term .Then the optimization problem is constructed by minimizing their joint reconstruction error. In order to achieve faster tracking multi-source targets.the computational complexity of the algorithm is further reduced by using the minimum error boundary constraint. The main contributions of this paper: 1) To establish an appearance model of joint sparse representation Infrared-visible; 2)To extend the L1-APG algorithm to multi-source target tracking; 3) To employ the least square error bound to multi-source target tracking; 4)To achieve the target updating.

The paper is organized as follows:In Section 2, the theory of accelerated proximal gradient is introduced; in Section 3, we present our tracking algorithm;in Section 4, To update target template is to achieve;in Section 5,experimental results and analysis are shown,we conclude in Section.

2. The Theory of APG

APG[15] is a kind of iterative search technology using the method of Nesterov. It is so simple and fast that target tracking meet with the real-time tracking. BAO et al.[16]improve the performance $\ell 1$ tracker using accelerated proximal gradient approach to achieve fast and robust target tracking. The APG algorithm is applied to solve the problem of convex optimization model as follows:

$$\min F(\alpha)=\min\{f(\alpha)+g(\alpha)\} \tag{1}$$

Where $f(\alpha)$ is a differentiable convex function, and $g(\alpha)$ is a non-smooth but convex function, $f(\alpha)$ and $g(\alpha)$ are separated by $F(\alpha)$. Actually, with the difference of the result of separating by $F(\alpha)$, Solving algorithms is estimated in different ways. In the paper, the function of $g(\alpha)$ equals $\lambda\|\alpha\|_1$ to layout from target tracking optimization function. So the APG algorithm in dealing with non-smooth convex problems and smooth problems have a very good performance. Ultimately, convex optimization model is established as follows:

$$\min F(\alpha)=\min\{f(\alpha)+\lambda\|\alpha\|_1\} \tag{2}$$

Assume that the function $f(\alpha)$ with Lipschitz continuous gradient and under the condition of $\|\nabla f(\alpha)-\nabla f(\alpha^*)\|\leq L(f)\|\alpha-\alpha^*\|$, where $L(f)>0$ is the constant of Lipschitz. In the k th iteration of traditional gradient method ,iterate α_{k-1} is update to:

$$\alpha_k=\alpha_{k-1}-t_k\nabla f(\alpha_{k-1}) \tag{3}$$

ISSN: 2688-8653

Here $t_k > 0$ represents a iterative step size. then aimed at $f(\alpha)$ in the point α_{k-1}, the second order approximation of the Taylor Series takes the second three terms in the series, We can be easily explained as the solution of a simple quadratic problem:

$$\alpha_k = \arg\min_{\alpha}\{f(\alpha_{k-1}) + \langle(\alpha-\alpha_{k-1}), \nabla f(\alpha_{k-1})\rangle + \frac{1}{2t_k}\|\alpha-\alpha_{k-1}\|_2^2\} \tag{4}$$

Here $<,>$ is the inner product the two matrixs. By neglecting constant terms $f(\alpha_{k-1})$ and $\langle(\alpha-\alpha_{k-1}, \nabla f(\alpha_{k-1}))\rangle$,(3)extend the gradient descent mechanism to solution(4),we gain

$$\alpha_k = \arg\min_{\alpha}\{\frac{1}{2t_k}\|\alpha-(\alpha_{k-1}-t_k\nabla f(\alpha_{k-1}))\|_2^2\} \tag{5}$$

The solution in the iterative(5) k th. it employs a convergence rate of $O(k^{-2})$, and satisfy the confidition of $F(\alpha_k) - \inf F(\alpha) \le O(\frac{L}{k^2})$,here L is Lipschitz contant. $F(\alpha_k)$ is the worth of $F(\alpha)$ in the interative k th and $\inf F(\alpha)$ is the low bound $F(\alpha)$. In order to design more accurate objective function,we may be set the iterative size $t_k = L^{-1}$, and combine the problem of(2)to achieve in the paper convex optimization model:

$$\alpha_k = \arg\min_{\alpha}\{\frac{L}{2}\left\|\alpha-(\alpha_{k-1}-\frac{1}{L}\nabla f(\alpha_{k-1}))\right\|_2^2 + \lambda\|\alpha\|_1\} \tag{6}$$

The solution(6)is a strongly convex function, therefore, there is only α to calculate the minimum of $F(\alpha)$. In sum, in the process of achieving the algorithm.we has consisted of three components in the paper: 1)Note that the function $g(\alpha) = \lambda\|\alpha\|_1$ is a non-smooth but convex function; 2)Suppose that the function $f(\alpha)$, with Lipschitz continuously differentiable gradient, is a soomth- convex function; 3)Assume that the iterative step size $t_k = L^{-1}$ to satisfy the precise of algorithm.

3. Proposed New Approach

3.1. Infrared-visible Images Using Joint Sparse Representation

In target tracking,the main thought of sparse representation [16]:appearance model, under different light and mutil-perspective orientation can belinearly sparse reconstructed by the over-complete dictionary of basis vectors as much as possible. Assume that a target template $T_I = [f_1,\ldots,f_n] \in R^{d\times n}$ in infrared images, the noisy template sets $E_I = [I,-I] \in R^{d\times 2d}$, $A_I = [T_I, E_I]$ is a infrared dictionary, $r_I \in R^d$ is the test datum, Then it can be defined as the solution of the following optimization problem in infrared images:

$$\arg\min_{\alpha_I}\{\|A_I\alpha_I - r_I\|_2^2 + \lambda\|\alpha_I\|_1\} \tag{7}$$

Here λ is a regularization parameter, Similarly, Suppose that a target template $T_V = [f_1,\ldots,f_n] \in R^{d\times n}$ in visible images, the nosiy template sets $E_V = [I,-I] \in R^{d\times 2d}$, $A_V = [T_V, E_V]$ is a visible dictionary, $r_V \in R^d$ is the test datum, Then it can be defined as the solution of the following optimization problem in visible images:

$$\arg\min_{\alpha_V}\{\|A_V\alpha_V - r_V\|_2^2 + \lambda\|\alpha_V\|_1\} \tag{8}$$

We can take full advantage of the complementary features of infrared and visible, Combining (7) and (8) together to joint sparse represent and expressing as follows

$$\arg\min_{\alpha_I,\alpha_V}\{\frac{L}{2}(\|A_I\alpha_I - r_I\|_2^2 + \|A_V\alpha_V - r_V\|_2^2) + \lambda(\|\alpha_I\|_1 + \|\alpha_V\|_1)\} \tag{9}$$

According to the equation, the most stable candidate can be selected α_I and α_V by minimizing their joint reconstruction error and the objective function coefficients are reconstructed mostly sparse. Equation (9) shows that whether it is determined that the state of candidate is pros or cons is minimized.by the sum of infrared reconstruction error and visible courtpart.

3.2. Realtime L1 Fusion Tracker Algorithm Using APG

Equation (9) belongs to solve 1-norm problem, the paper ,different form the solution of soving 2-norm in reference [11], employs APG algorithm introducted in Section 2, Reference[15] has proved APG algorithm in theory, which is more faster than the conventional gradient descent method. so it can obtain the optimal sparse factors with the more quick speed. the equation (9) is derived and solved as follows by combining equations (6),

$$\begin{aligned}\langle\alpha_I^*,\alpha_V^*\rangle &= \arg\min_{\alpha_I,\alpha_V}\{\frac{L}{2}(\|A_I\alpha_I - r_I\|_2^2 + \|A_V\alpha_V - r_V\|_2^2) + \lambda(\|\alpha_I\|_1 + \|\alpha_V\|_1)\}\\ &= \arg\min_{\alpha_I,\alpha_V}\{(\frac{L}{2}\left\|\alpha_I - (\alpha_I^{k-1} - \frac{1}{L}\nabla f(\alpha_I^{k-1}))\right\|_2^2 + \lambda\|\alpha_I\|_1) + (\frac{L}{2}\left\|\alpha_V - (\alpha_V^{k-1} - \frac{1}{L}\nabla f(\alpha_V^{k-1}))\right\|_2^2 + \lambda\|\alpha_V\|_1)\}\\ &= \arg\min_{\alpha_I,\alpha_V}\{(\frac{L}{2}\|\alpha_I - c_I^k\|_2^2 + \lambda\|\alpha_I\|_1) + \frac{L}{2}(\|\alpha_V - c_V^k\|_2^2 + \lambda\|\alpha_V\|_1)\}\end{aligned} \tag{10}$$

Here, $c_I^k = \alpha_I^{k-1} - t_k\nabla f(\alpha_I^{k-1}) = \alpha_I^{k-1} - 2t_kA^T(A\alpha_I^{k-1} - r_I)$, $c_V^k = \alpha_V^{k-1} - t_k\nabla f(\alpha_V^{k-1}) = \alpha_V^{k-1} - 2t_kA^T(A\alpha_V^{k-1} - r_V)$ The convergence of L1-APG can be achieved $O(k^{-2})$. For the equation(9)the soving real-time numerical APG algorithm can be summarized in algorithm.

3.3. Reducing the Number of Calculating L1-APG

Even though APG algorithm was proposed to decrease the compution time for L1 tracker, the computational complexity is still very high with a large number of particles, which is very low. A minimal error bounding approach is provided in [17]to reduce the amount of needed 1-norm minimizations.the smaller weight particles may be discarded in resampling process, our method is based on the following observation:

$$p(z_k \mid x_k) = \exp\{-\eta(\|A_I\alpha_I^* - r_I\|^2 + \|A_V\alpha_V^* - r_V\|^2)\} \tag{11}$$

here η is the constraint of the Gaussian kernel function. There is natural lower bound for reconstruction error.

$$\begin{cases}\left\|A_I\alpha_I^*-r_I\right\|_2^2\geq\left\|A_I\hat{\alpha}_I-r_I\right\|_2^2\\ \hat{\alpha}_I=\arg\min_{\alpha_I}\left\|A_I\alpha_I-r_I\right\|_2^2\end{cases},\begin{cases}\left\|A_V\alpha_V^*-r_V\right\|_2^2\geq\left\|A_V\hat{\alpha}_V-r_V\right\|_2^2\\ \hat{\alpha}_V=\arg\min_{\alpha_V}\left\|A_V\alpha_V-r_V\right\|_2^2\end{cases} \tag{12}$$

Where $\min\left\|A_I\alpha_I^*-r_I\right\|_2^2$ and $\min\left\|A_V\alpha_V^*-r_V\right\|_2^2$ are minimal error bound, Similarly, for the observation like hood function we immediately have

$$p(z_k|x_k)\leq q(z_k|x_k)=\exp\{-\eta(\left\|A_I\hat{\alpha}_I-r_I\right\|^2+\left\|A_V\hat{\alpha}_V-r_V\right\|^2)\} \tag{13}$$

Here $q(z_k\mid x_k)$ is derived the probability of its upper bound. Therefore, if the sample x_k will not appear with being process, its posterior probability meet the condition of $p(z_k|x_k)\leq q(z_k|x_k)<\frac{1}{2N}\sum_{j=1}^{i-1}p(z_k|x_k^j)$ the sample x_k will be discarded, i, j is the number of practiles. Note that the total number of particles is n, Then the overall time of the algorithm should be less than $nO(k^{-2})$. Thus the algorithm can effectively redue the times of calculating L1-APG, at the same time,and ensure the accuracy and real-time.

4. Experiment and Analysis

To test the effectiveness of the method, four sets of image sequences were chosen to test JSRT performance which compare with $\ell1$ tracker ($\ell1$ tracker, L1T) [9], Fuzzified Region Dynamic fusion, FRD tracker [18] algorithm in the respect of qualitative and quantitative analysis.these methods here tested was originate from original author offering. In the experiment, the target was initialized with a hand marked rectangle region in the first frame, the experimental environment for the Intel (R) Core (TM) i5-3470 CPU @ 3.20 GHZ, 8.00GB memory of the PC, the algorithm was achieved by Matlab R2014a. In this paper, the experimental parameters that we set Lipschitz constant is 8, τ=0.01.

Quantitative Comparison: There are four evaluation criterias in quantitative comparison in this paper,its includes average error $\bar{\phi}=1/\sum_{i=1}^{n}\phi_i$,the position error $\phi=\sqrt{(x_G-x_T)^2+(y_G-y_T)^2}$, overlap rate $\xi=area\,|\gamma_G\cap\gamma_T\,|\,/\,area\,|\gamma_G\cup\gamma_T\,|$, the average rate of the overlap $\bar{\xi}=1/\sum_{i=1}^{n}\xi_i$ and success rate sr , the central position of the real target is defined (x_T,y_T),(x_G,y_G) is the central position of tracking result, γ_T denotes the ground truth region designed by manully selecting as the best match of the target, γ_G denotes the target window provided by the tracker,where $\cap$ and $\cup$ represent the intersection and union of two regions,respectively, and $|\cdot|$ denotes the number of pixels in the area,The ratios of successful frames more than a half of all the frames tested, Table 1 presents the results of quantiative comparison. It is no diffierence to find that compared with FRD, L1TVS, L1TIR algorithm,our approach can obtain a better overall performance.

In order to measure the real time of this paper,video1 is used to test the time complexity and success rate of result tracking. Under the same number of particles, the success rate of JSRT is obviously higher than L1TVS and L1TIR trackers,which with the increase of the number of particles is particularly evident. when the number of particles decreasing, nevertheless ,the success rates of L1TVS and L1TIR are particularly went down. For instance, the number of particles in JSRT is 100,whose rate is basically equal to that in L1TVS tracker the number of particles is 600 .the experiment result indicate that fusion tracking Infrared-visible target using L1-APG has apparently great advantage in real time.

Table 1. Quantitative results of the tracking methods

method	video1(320×240)			video2(320×240)			video3(320×240)			video4(480×360)		
	$\overline{\varphi}$	$\overline{\xi}$	sr	$\overline{\varphi}$	$\overline{\xi}$	sr	$\overline{\varphi}$	$\overline{\xi}$	sr	$\overline{\varphi}$	$\overline{\xi}$	sr
JSRT	7.140	0.669	0.955	2.922	0.802	1.000	2.363	0.851	1.000	16.94	0.690	0.944
L1TIR	15.425	0.456	0.690	3.414	0.800	1.000	2.076	0.856	1.000	17.02	0.675	0.968
L1TVS	14.205	0.533	0.800	2.361	0.825	1.000	4.488	0.726	0.833	16.15	0.529	0.480
FRD	20.422	0.362	0.548	3.103	0.811	1.000	3.980	0.795	0.910	203.0	0.129	0.128

5. Conclusion

The paper,introduced fusion tracking sparse representation,is established the appearance of Infrared-visible joint sparse representation model, by which joint reconstruction error the objective optimization problem is built.APG method is used to solve L1 problem and then the least square error is employed to reduce particle to resampling number,which enjoys the less time complexity of the algorithm,and achieves real-time joint tracking.our paper also for peer made an reference.Taking into account the different circumstances advantages in infrared and visible senior, Infrared-visible target model with adaptive weights updating mechanisms will be proposed to achieve more robust tracking in the future research.

Acknowledgments

This paper was financially supported by general projects in Natural Science held by Guangdong University of Science and Technology NO.GKY-2018KYYB-11, NO. GKY-2019KYYB-31, 36, 42 and NO.GKY-2020KYYBK-22, 24, 27, General projects of social science and technology development in Dongguan in 2020NO.2020507154643.

References

[1] Wu Y, Lim J, Yang M H. Online Object Tracking: A Benchmark[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2013:2411-2418.

[2] Nguyen A H , Mai L , Do H N . An Implementation of a Robust Visual Object Tracking System[C]// 2020 International Conference on Advanced Technologies for Communications (ATC). US, 2020.

[3] Kracunov M , Bastica M , Tesovic J . Object tracking in video signals using Compressive Sensing[J]. 2019.

[4] Leichter I. Mean Shift Trackers with Cross-Bin Metrics[J]. Pattern Analysis & Machine Intelligence IEEE Transactions on, 2012, 34(4):695-706.

[5] ZHANG C L,TANG Y P,LI Z X, et al. Dual-kernel tracking approach based on second-order spatiogram [J]. Journal of Electronics & Information Technology, 2015, 37(7):1660-1666.

[6] Zhang S, Sui Y, Yu X, et al. Hybrid support vector machines for robust object tracking[J]. Pattern Recognition, 2015, 48(8):2474-2488.

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

[8] Liu Zhe, Chen Ken, Zheng Ziwei. Object Tracking Algorithm Based on HOG and Multiple-instance Online Learning[J].Computer Engineering,2015,41(1):158-163.

[9] JING L. Incremental learning for robust visual tracking[J]. International Journal of Computer Vision, 2015, 77(1-3):125-141.

[10] MEI X, LING H. Robust visual tracking using ℓ1 minimization[C]// Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009:1436-1443.

[11] SUN F, LIU H. Fusion tracking in color and infrared images using joint sparse representation[J]. Science China Information Sciences, 2012, 55(3):590-599.

[12] CHENG Y M,WANG J X,WEI K, et al. Visible/Infrared dual-channel target tracking based on weighted mean-shift[J]. ACTA PH OTONICA SINICA, 2010, 39(2):357-363.

[13] ZHAO G P,BO Y M,CHEN Y. A method for tracking object in infrared and visible image based on multiple features[J].ACTA ARMAMENTARII, 2011, 32(4):445-451.

[14] YUN X, JING Z, XIAO G, et al. A compressive tracking based on time-space Kalman fusion model[J]. Science China Information Sciences, 2016(1):1-15.

[15] TSENG P. On accelerated proximal gradient methods for convex-concave optimization[J]. Siam Journal on Optimization, 2008.

[16] BAO C. Real time robust ℓ1 tracker using accelerated proximal gradient approach[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2012:1830-1837.

[17] MEI X, LING H, WU Y, et al. Minimum error bounded efficient ℓ1 tracker with occlusion detection [C]//.Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011: 1257-1264.

[18] Xiao X , Xie J , Niu J , et al. A Novel Image Fusion Method for Water Body Extraction Based on Optimal Band Combination[J]. Traitement du Signal, 2020, 37(2):195-207.