Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 7, July 2016, pg.283 – 289

Implementation of Robust Visual Tracker by using Weighted Spatio-Temporal Context Learning Algorithm

Ms. Pallavi Talele[1], Prof. Dinesh D. Patil[2]

[1]PG Student, Department of Computer Science and Engineering, Shri Sant Gadge Baba College of Engineering and Technology, Bhusawal, Maharashtra, India

[2]Associate Professor and Head, Department of Computer Science and Engineering, Shri Sant Gadge Baba College of Engineering and Technology, Bhusawal, Maharashtra, India

[1] talele.pallavi1205@gmail.com; [2] dineshonly@gmail.com

Abstract— In computer vision object tracking is an area that has been explored but still there remain scope for improvement and innovation. To design a robust visual tracker itself is a challenging task apart from which its role in forensics and videos surveillance is significant. In this paper we have implemented robust visual tracker using weighted spatio-temporal learning algorithm. Visual tracker itself is applicable to both image sequence and video; hence in our work we have used both frames generated from videos as image sequence and avi format files for video purposes. Both tracking is done separately. Dataset and videos chosen are from standard dataset and particularly having illumination challenges so that robust nature of algorithm is proven. Experimental results show that the tracking done using WSTC algorithm is better as compared to existing systems.

Keywords— Visual Tracking, Context learning, Computer Vision, Pattern Matching, Learning

---

I. INTRODUCTION

Visual tracking is a future of research in imaging due to its wide range of applications such as motion analysis, activity recognition, surveillance and human computer interaction. Visual tracking is one of the most active research topics due to its wide range of applications such as motion analysis, activity recognition, surveillance, and human computer interaction. The main challenge for robust visual tracking is to handle large appearance changes of the target object and the background over time due to occlusion, illumination changes, and pose variation. Tracking algorithms can be classified as either discriminative or generative methods, which makes trade-offs between effectiveness and efficiency of an appearance model. The tracking problem is formulated by computing a confidence map, spatial context model and context prior model. Human visual system that exploits context to help resolving ambiguities in complex scenes efficiently and effectively is the first stage in algorithmic model to compute the spatio-temporal relationships between the object and its local contexts. To design robust visual tracker is a challenging problem, there are many factors such as illumination changes, appearance changes, rotation, partial or full occlusions which causes disturbance in accuracy. Among existing trackers, correlation filter based tracker is a fast and robust method with provides resistance earlier mentioned factors. Motivated by spatio-temporal context learning algorithm STC as it treats the whole region of the context equally, which weakens the effectiveness of the context information. We have proposed a novel weighted spatio-temporal context (WSTC) learning algorithm

II. RELATED WORK

Visual tracking is one of the most active research topics due to its wide range of applications such as motion analysis, activity recognition, surveillance, and human-computer interaction, to name a few [10]. In today's real world appearances of the target and its surroundings change continuously this provides effective information to track the target robustly. However, enough attention has not been paid to the spatio-temporal appearance information as a result authors in [11] proposed robust spatio-temporal context model based tracker to complete the tracking task in unconstrained environments. This tracker is constructed with temporal and spatial appearance context models. The historical appearance of the target to prevent the tracker from drifting to the background in a long-term tracking is captured by temporal appearance context model. However spatial appearance context model integrates contributors to build a supporting field. The contributors are the patches with the same size of the target at the key-points automatically discovered around the target. The robustness of the tracker in complex environments is ensured by constructive supporting field that provides much more information than the appearance of the target itself.

In [12], an Incremental Visual Tracker (IVT) is proposed, which adaptively updates its subspace-based appearance model with the sequential appearance variations. Fragment-based tracker [13] describes the target with multiple local patch histograms, which integrates the inner structure of the target and handles partial occlusion very well. Avidan [14] integrates the Support Vector Machine (SVM) classifier into the optical flow framework for car tracking. Grabner et al. propose an efficient supervised online boosting tracking method [15]. A semi-supervised version [16] is proposed, in which the labelled data in the first frame is used whereas subsequent training samples are left unlabeled. Bakenko et al. [17] use Multiple Instance Learning (MIL) to handle the unreliable labelled positive and negative data obtained online to mitigate the drift problem. In all these trackers, only the appearance of target is considered, but the relationships between target and its background are not fully exploited. To track the objects, the detection responses at different frames are linked together to form the object trajectories. Wu and Nevatia [19] define an affinity measure between detection responses based on cues from position, size, and color and use the Hungarian algorithm [20] to associate object hypotheses and detection responses. New object trajectories are initialized whenever the detection responses do not match with any existing trajectories for a certain number of frames; old trajectories are terminated when they are lost by the detector for a certain number of frames. Li et al. [21] and Okuma et al. [22] use particle filter methods to associate the detection responses of an unknown number of objects.

III.PROPOSED SYSTEM

Designing a robust visual tracker is itself a challenging problem when that associated with illumination changes, appearance changes and also rotation, partial changes and much more. In this section we explained the proposed system workflow along with few related terms.

A. Problem formulation

The problem is implementation of robust visual tracker means such tracker that can be useful under extreme challenging conditions for computer vision; the tracker is applicable in video as well as image sequences. We implement proposed system using weighted spatio-temporal context learning algorithm.

B. Proposed system workflow

Step 1. *Start*
Step 2. *Select whether to load image sequence or load video*
Step 3. *For image sequence*
- a. *Select the option load image sequence*
- b. *The tracking process start*
    - i. *First initialize the boundary of object to be tracked*
    - ii. *Set boundaries and build confidence map*
    - iii. *Go for get context so to set boundaries and normalize with weighted map*
    - iv. *Use spatio-temporal context model and build visualization*
    - v. *Position rectangle and build colormap with ouput*

Step 4. *For load video part*
- a. *Select option load video*
- b. *Choose video we consider only avi files.*
- c. *We use optical flow method for tracking*
    - i. *Read a file and in its binary form*

ii. *Convert file into RGB format so video frames generated are used.*
iii. *Build system object Blob analysis for gaps removal in neighbourhood*
iv. *Next use Morphological erosion to thin out portions*
v. *Write out how many object in given case cars are tracked down.*

Step 5. *Select exit option to quit application*
Step 6. *Continue with above steps for different videos to be tested.*

IV. EXPERIMENTAL RESULTS

The implementation of proposed work is done using MATLAB, for conducting experiments we had used the standard dataset available which comprises of videos and sequences that where having different illumination challenges, having variation in contrast, background and blurring, so that for tracking of object subject to different challenges. For illustrations we have taken two separate options in our simple GUI that is shown in figure 1.

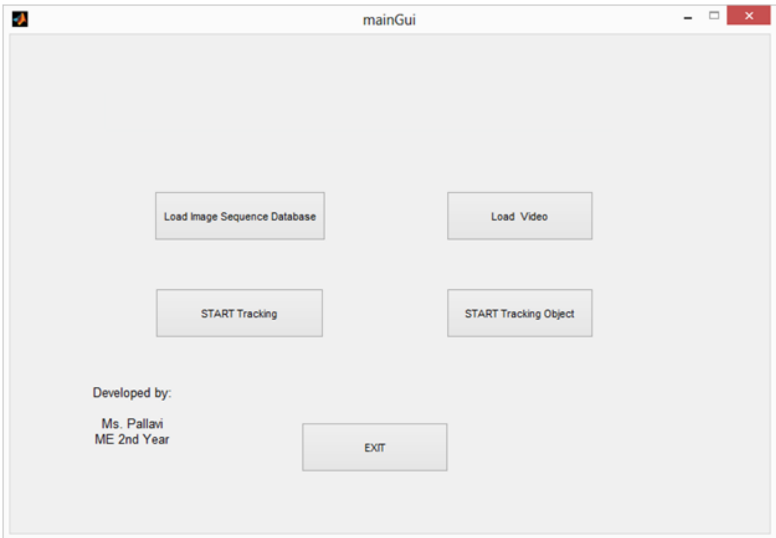


Fig. 1 Main GUI of proposed system

For tracking we have suggest two options to load image sequence and second to load video, the different types of videos used sample is as shown in table 1.

TABLE I
DESCRIPITON ABOUT VIDEOS USED FOR TRACKING

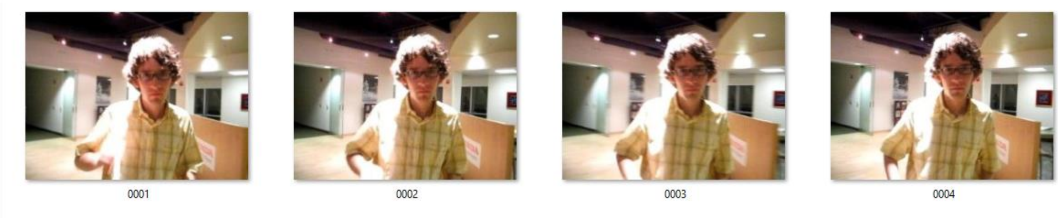
Sr. No.	Video filename	Bit rate	Dimensions in pixels	No. of frames generated for image sequence
1	car_traffic.avi	238 kbps	160 x 120	120
2.	street.avi	240 kbps	320 x 240	390
3.	highway.avi	240 kbps	320 x 240	540

For the tom_miker.avi video we have about 800 frames generated which are having illumination challenges out of these for tracking of image sequence that his face figure 2(a) – (f) shows sample image frame sequence used in tracking of object.

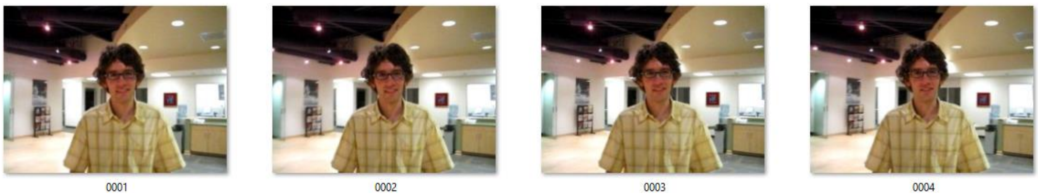
(a)

(b)

(c)



(d)

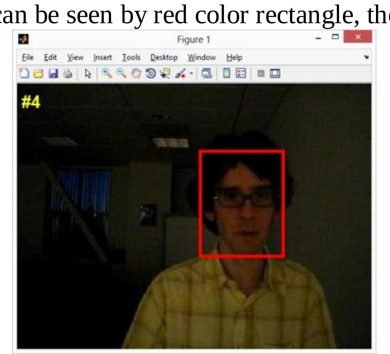


(e)

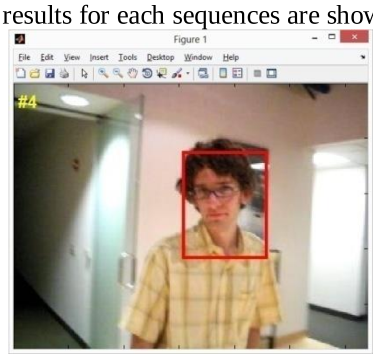
(f)

Fig. 2 (a)-(f) Image Frame sequence for tracking having different type of illumination challenges.

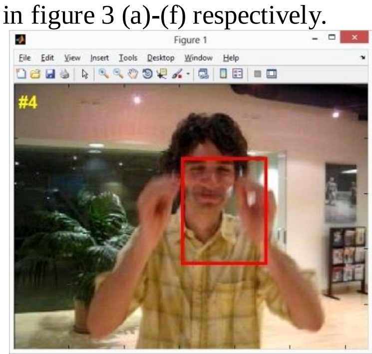
For tracking we implemented tracking of face in each of frame sequence mentioned, therefore the tracked object can be seen by red color rectangle, the results for each sequences are shown in figure 3 (a)-(f) respectively.



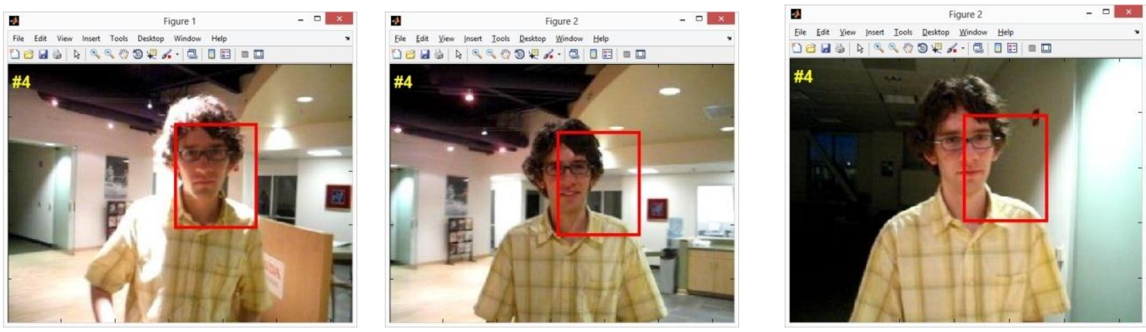
(a)



(b)



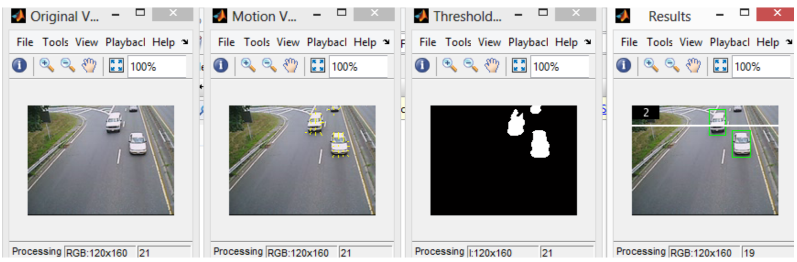
(c)



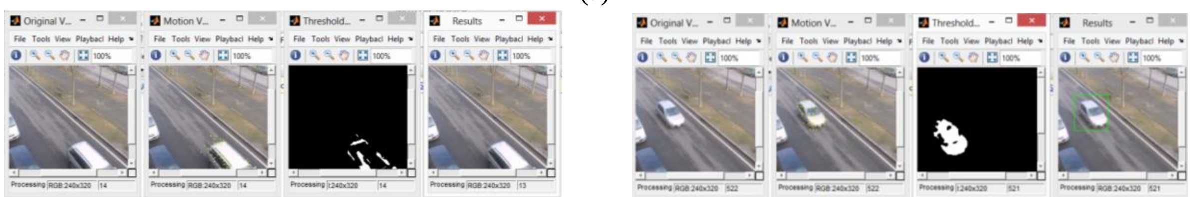
(d) (e) (f)

Fig. 3 Results of tracking in image frame sequences for tim_miker.avi image sequence.

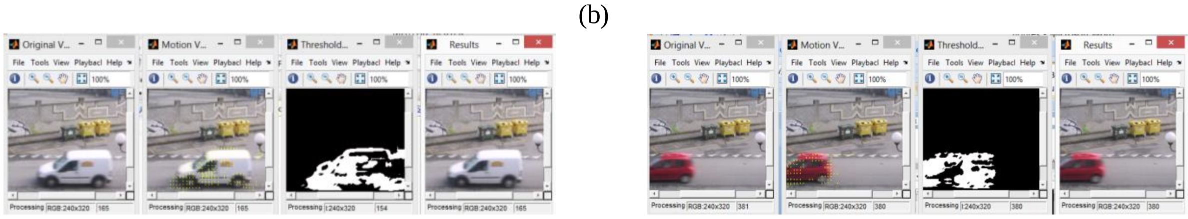
Next part is about tracking in for video, after clicking on load video option, after background intermediate steps we have the generated result sequences as shown below in figure 4 for videos tabulated in table 1.



(a)



(b)



(c)

Fig. 4 Results of tracking in video using proposed method for (a) car_traffic.avi (b) highway.avi (c) street.avi

V. Conclusions

In this paper we have implemented a robust visual tracker using weighted spatio-temporal context learning algorithm, in which we have illustrated by performing experiments on image frame sequences as well as video file of avi format. Algorithm considers the surrounding context discriminatively and generates a weighted map by evaluating the importance of different region. According to the consistency between the movement of the object and other regions, the surrounding context could be divided into three classes approximately. Since using learning and use of weighted map, the results from tracking are more accurate for different types of illumination changes. Results obtained shows better performance than that of existing algorithms.

Acknowledgement

We would like to articulate our profound gratitude and indebtedness to those persons who helped us in completion of the paper. First and foremost, we would like to convey our gratefulness our Principal Dr. R. P. Singh for his continuing support. We would like to dedicate this work our parents and would like to thank all the staff members from Department of Computer Science and Engineering, SSGBCOET Bhusawal Maharashtra. Also our special thanks to all the writers of reference paper that are referred by us.

REFERENCES

[1] Wei Zhong, Huchuan Lu, Ming-Hsuan Yang, "Robust Object Tracking via Sparsity-based Collaborative Model", In CVPR 2012.

[2] J. Wright, Y. Ma, J. Maral, G. Sapiro, T. Huang, and S. Yan. Sparse representation for computer vision and pattern recognition. Proceedings of the IEEE, 98(6):1031–1044, 2010.

[3] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. PAMI, 31(2):210–227, 2009.

[4] J. Yang, J. Wright, T. S. Huang, and Y. Ma. Image super-resolution via sparse representation. TIP, 19(11):2861–2873, 2010.

[5] X. Mei and H. Ling. Robust visual tracking using ℓ1 minimization. In ICCV, 2009.

[6] B. Liu, J. Huang, L. Yang, and C. Kulikowsk. Robust tracking using local sparse appearance model and k-selection. In CVPR, 2011.

[7] B. Liu, L. Yang, J. Huang, P. Meer, L. Gong, and C. Kulikowski. Robust and fast collaborative tracking with two stage sparse optimization. In ECCV, 2010.

[8] Adam, E. Rivlin, and I. Shimshoni. Robust fragments-based tracking using the integral histogram. In CVPR, 2006.

[9] Kaihua Zhang, Lei Zhang, Qingshan Liu, David Zhang, and Ming-Hsuan Yang, "Fast Visual Tracking via Dense Spatio-Temporal Context Learning", In ECCV 2014, pp. 1-15

[10] Yilmaz, A., Javed, O., Shah, M.: Object tracking: A survey. ACM Computing Surveys 38(4) (2006)

[11] Longyin Wen, Zhaowei Cai, Zhen Lei, Member, IEEE, Dong Yi, and Stan Z. Li, Fellow, IEEE, "Robust Online Learned Spatio-Temporal Context Model for Visual Tracking", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 23, NO. 2, FEBRUARY 2014, pp. 785-796.

[12] J. Lim, D. A. Ross, R.-S. Lin, and M.-H. Yang, "Incremental learning for visual tracking," in Proc. NIPS, 2004, pp. 1–8.

[13] Adam, E. Rivlin, and I. Shimshoni, "Robust fragments-based tracking using the integral histogram," in Proc. IEEE CVPR, Jun. 2006, pp. 798–805.

[14] S. Avidan, "Support vector tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 8, pp. 1064–1072, Aug. 2004.

[15] H. Grabner and H. Bischof, "On-line boosting and vision," in Proc. IEEE CVPR, Jun. 2006, pp. 260–267.

[16] H. Grabner, C. Leistner, and H. Bischof, "Semi-supervised on-line boosting for robust tracking," in Proc. 10th ECCV, Oct. 2008, pp. 234–247.

[17] B. Babenko, M.-H. Yang, and S. J. Belongie, "Visual tracking with online multiple instance learning," in Proc. CVPR, 2009, pp. 983–990.

[18] Bo Wu, Li Zhang, Vivek Kumar Singh, and Ram Nevatia, "Robust Object Tracking based on Detection with Soft Decision", 2008, pp. 1-8.

[19] B. Wu, and R. Nevatia. Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors. International Journal of Computer Vision, 2007.

[20] H.W. Kuhn. The Hungarian method for the assignment problem. Naval Research Logistics Quarterly, 2: 83-87, 1955.

[21] Y. Li, H. Ai, T. Yamashita, S. Lao, and M. Kawade. Tracking in Low Frame Rate Video: A Cascade Particle Filter with Discriminative Observers of Different Lifespans. CVPR 2007.

[22] K. Okuma, A. Taleghani, N. D. Freitas, J. J. Little, and D. G. Lowe. A Boosted Particle Filter: Multitarget Detection and Tracking. ECCV 2004.

[23] M. Li, W. Chen, K. Huang, and T. Tan, "Visual tracking via incremental self-tuning particle filtering on the affine group," in Proc. IEEE CVPR, Jun. 2010, pp. 1315–1322.

[24] X. Zhang, W. Hu, S. J. Maybank, X. Li, and M. Zhu, "Sequential particle swarm optimization for visual tracking," in Proc. IEEE CVPR, Jun. 2008, pp. 1–8.

[25] M. Kristan, J. Pers, S. Kovacic, and A. Leonardis, "A local-motion-based probabilistic model for visual tracking," Pattern Recognit., vol. 42, no. 9, pp. 2160–2168, Sep. 2009.

[26] J.-Y. Lu, Y.-C. Wei, and A. C.-W. Tang, "Visual tracking using compensated motion model for mobile cameras," in Proc. IEEE ICIP, Sep. 2011, pp. 489–492.

[27] David S Bolme, Bruce A Draper, and J Ross Beveridge, "Average of synthetic exact filters," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009, pp. 2105–2112.

[28] Abhijit Mahalanobis, BVK Vijaya Kumar, Sewoons Song, SRF Sims, and JF Epperson, "Unconstrained correlation filters," Applied Optics, vol. 33, no. 17, pp. 3751–3759, 1994.

[29] David S Bolme, J Ross Beveridge, Bruce A Draper, and Yui Man Lui, "Visual object tracking using adaptive correlation filters," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010, pp. 2544–2550.

[30] Kaihua Zhang, Lei Zhang, Ming-Hsuan Yang, and David Zhang, “Fast tracking via spatio-temporal context learning,” arXiv preprint arXiv:1311.1939, 2013.
[31] Zhongqian Sun, Hongxun Yao, Shengping Zhang, and Xin Sun, “Robust visual tracking via context objects computing,” in Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE, 2011, pp.509–512.
[32] Longyin Wen, Zhaowei Cai, Zhen Lei, Dong Yi, and Stan Z Li, “Online spatio-temporal structural context learning for visual tracking,” in Computer Vision–ECCV 2012, pp. 716–729. Springer, 2012.
[33] Helmut Grabner, Jiri Matas, Luc Van Gool, and Philippe Cattin, “Tracking the invisible: Learning where the object might be,” in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010,pp. 1285–1292.

