IJARIIT INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 1)
Available online at www.ijariit.com

Survey of Generative and Discriminative Appearance Models in Visual Object Tracking

Remya K. V
remyavgd@gmail.com
Cochin College of Engineering and Technology, Valanchery, Kerala

Vipin Krishnan C. V
vipinkrishnanm@gmail.com
Cochin College of Engineering and Technology, Valancheri, Malappuram, Kerala

ABSTRACT

Visual object tracking is a challenging task in computer vision applications. The basic statistical appearance modeling techniques are discriminative and generative. In both cases, online learning is v essential to nullify the error due to large pose changes, illumination variations and appearance changes of the tracking framework. This paper briefly introduces the challenges and applications of visual tracking and focuses on discussing the state-of-the-art online-learning based tracking methods by category. In this paper, the existing statistical schemes for tracking-by-detection are reviewed according to their appearance model creation mechanism; generative and discriminative.

Keywords: *Visual Object Tracking, Modeling Techniques, Online Matric Learning.*

I INTRODUCTION

Visual tracking is defined as the problem of finding the motion of a target given a sequence of images based on different frames in a video. Computer technology advances lead visual tracking to become one of the most popular topics in the field of computer vision. Generally, the visual object tracking system is composed of four different modules: object initialization, appearance modeling, motion estimation, and object localization. So many challenges are present in the object tracking criteria. A large variety of methods based on online learning are proposed so as to nullify the effect of errors in tracking. Even though there are so many challenges in results as the application becomes more complicated. Some of the challenges of the visual object tracking are illumination changes, dynamic background, occlusion, clutter, camouflage, the presence of shadows, and appearance changes etc.

Traditional tracking methods use fixed modelsino the starting stage of tracking task. It normally fails due to the inevitable appearance changes.These are not only from the object itself, but also from surrounding environment, such as varying changes, camera motion , camera scale, and occlusion. Basic object tracking scenario deals with observation model, integration of observation, target detection along with update model as in the figure.1the basic statistical models are again subdivided based on the design of appearance of objects

In this paper, some of the visual objects tracking methods are categorized based on the statistical method used for the appearance model adopted. There is a broad range of applications of object tracking that motivate the interests of researchers worldwide. Video surveillance is a very popular one. Surveillance systems are not only for recording the observed visual information, but also extracting motion information and, more recently, to analyze suspiciously

Behaviors in the scene. One can visually track airplanes, vehicles, animals, micro-organisms or other moving objects, but detecting and tracking people is of great interest. For instance, vision-based people-counting applications can provide important information for public transport, traffic congestion, tourism, and retail and security tasks. Tracking humans is also an important step for human-computer interaction (HCI).

Some of the constraints that generally imposed during object tracking are:

- Object motion is smooth with no abrupt changes
- No sudden changes in the background
- Gradual changes in the appearance of object
- Fixed camera
- Number and size of objects
- Limited amount of occlusion

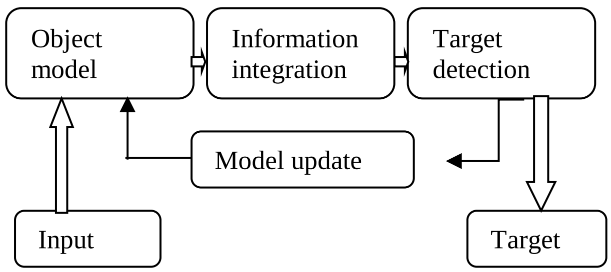

Figure 1.Flowchart of Visual Tracking

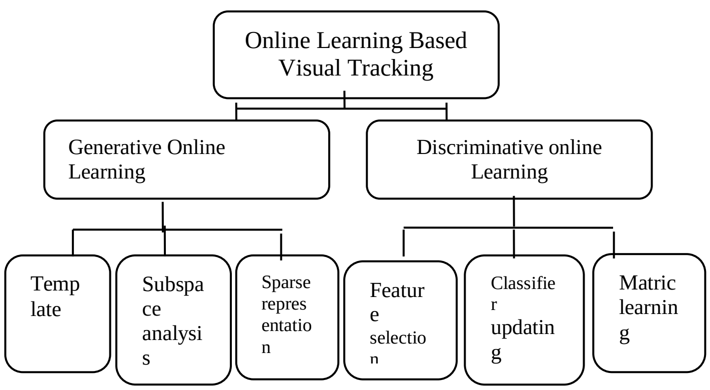

Figure 2: Methodologies in the Statistical Appearance Models

2 GENERATIVE LEARNING METHODS

The generative appearance models mainly concentrate on how to accurately fit the data from the object class. However, it is very difficult to verify the correctness of the specified model in practice. By introducing online-update mechanisms, they incrementally learn visual representation for the foreground object region information while ignoring the influence of the background. Traditional online learning methods are adopted to track an object by searching for regions most similar to the target model. The online learning strategy is embedded in the tracking framework to update the appearance model of the target adaptively in response to appearance variations. Different methods under generative model are described below.

2.1methods Based on Templates

Suitable appearance model is needed to avoid the tracking errors. Traditional template-matching methods do not perform well due to fixed models. Hence dynamic templates based on online learning are used. Online learning will improve the performance by making tracker to more adaptive for appearance changes. The online method uses the position, size or other details of the target in the first frame and is used to tracking the object from a subsequent frame.

Template based generative appearance model is discussed in [1].this method used the online expectation-maximization (EM) algorithms to update the appearance model and derive the tracking parameters efficiently for motion based trackers. The purpose of online EM algorithm to estimate appearance model parameters and update the histogram space over time. Such an algorithm perfectly applied in [2].these methods are more simple and stable. Even though there is more iterations present and may lead to tracking task failure.

In order to eliminate the impact of the background on the template and reduce the computation scale of matching during tracking, a new object tracking algorithm is proposed in [3] based on dynamic template and motion prediction. When using the similarity sequential detection algorithm (SSDA), it matches only with the pixels in the template, which reduces the risk of disturbances from the background. Moreover, the template is updated continuously to ensure the correctness of the template data. By taking the area of the actual curve in a region as the background value and using the least square of the difference between the simulation sequence and the first order accumulating generated operator (AGO) sequence.

To overcome the limitation of the deterministic optimization approach, in the paper [3] present a novel particle filtering approach to template-based visual tracking. We formulate the problem as a particle filtering problem on matrix Lie groups, specifically the three-dimensional Special Linear group and the two-dimensional affine group

Template based methods use fixed number of basic trackers. Hence during the complicated videos where severe changes occur between the frames, The tracking may fail.to solve this problem a new method of visual tracking decomposition[4] is presented.

2.2 Methods Based on Subspace Analysis

Instead of using simple templates to represent the appearance model for tracking, an online learned subspace representation can be used to provide a compact representation of a target and reflect appearance changes during tracking.

In paper [5], present an adaptive probabilistic tracking algorithm that updates the models using an incremental update of Eigen basis. To track objects in two views, we use an effective probabilistic method for sampling affine motion parameters with priors and predicting its location with a maximum *a* posteriori estimate. An incremental mean update was added to learning algorithm [6], thus a low-dimensional Eigen space representation of the appearance of the target was learned in order to deal with changes.

An online incremental learning is proposed in [7]. This paper presents an effective online tensor subspace learning algorithm which models the appearance changes of a target by incrementally learning a low-order tensor Eigen space representation by adaptively updating the sample mean and Eigen basis. Tracking then is led by the state inference

Within the framework in which a particle filter is used for propagating sample distributions over the time. A novel likelihood function, based on the tensor reconstruction error norm is developed to measure the similarity between the test image and the learned tensor subspace model during the tracking.

2.3 Methods Based on Sparse Representation

The concept of sparse representation recently attracted the computer vision community due to its discriminative ability. Sparse representation has been applied to various computer vision tasks including face recognition, image video restoration, image denoising, action recognition, super resolution, and tracking the l_1 minimization tracker proposed by Mei et al. [8] uses the low resolution target image along with trivial templates as dictionary elements. The candidate patches can be represented as a sparse linear combination of the dictionary elements. To localize the object in the future frames, the authors use particle filter framework. Here, each particle is an image patch obtained from the spatial neighborhood of previous object center. The particle that minimizes the projection error indicates the location of object in the current frame. Typically, hundreds of particles are used for localizing the object. The performance of the tracker relies on the number of particles used. This tracker is computationally expensive and not suitable for real-time tracking.

A faster version of the above work [9] was proposed by Bao et al. [10], here the $l2$ norm regularization over trivial templates are added to the $l1$ minimization problem. However, our experiments show that the speed of the above algorithm is achieved at the cost of its tracking performance. Other faster version was proposed by Li et al. [11]. Jia et al. [12] used structural local sparse appearance model with an alignment-pooling method for tracking the object. All these sparse trackers are implemented in particle filter framework and runs at the low frame rate, hence not suitable for real-time tracking. A recent real-time tracker proposed by Zhang et al. [13] uses a sparse measurement matrix to reduce the dimension of foreground and background sample and uses naive Bayes classifier for classifying object from the background.

3. DISCRIMINATIVE ONLINE LEARNING METHODS

Discriminative methods formulate tracking as a classification problem. The trained classifier is used to discriminate the target from the background and can be updated online. This method utilizes the information about the target and background simultaneously. A binary classifier is trained to distinguish the target from the background and is updated online to handle appearance and environmental Changes.

3.1 Methods Based on Online Feature Selection

The discriminative ability of tracking is based on the feature space used. If the feature is very accurate then, a simple tracker can track the object. It's better to consider online feature selection compared with a fixed set of features for better performance. Hence a feature ranking mechanism is essential in these methodologies.

In the last decade, boosting based discriminative appearance models (BDAM) have been widely used in visual object tracking because of their powerful discriminative learning capabilities.BDAMS first train the classifier based on the previous frames and subsequently trained a classifier to evaluate possible object regions in the current frame.

To improve the feature selection process , Liu and Yu [2007] utilize gradient based feature selection to construct BDAM require an initial set of weak classifiers and it may lead to poor performance in general object tracking.it have less ability to capture correlation between the features.

For avoiding these problems a feature weighting strategy is provided to consider all the features from the feature pool. This weighting technique is better for object tracking.Avidan [2007] construct a confidence map by pixel classification using an ensemble of the online learned weak classifier, which is trained by a feature weighting based boosting approach.

3.2 Methods based on online updating classifier

If there is no prior knowledge about the object to be tracked then binary classifier needs to be trained online.in this method, only the initial position of the target is given in the first frame. Avidan [14] uses AdBoost.The strong classifier is then used to label pixels in the next frame as either belonging to the object or the background, giving a confidence map. The peak of the map, and hence the new position of the object, is found using mean shift. The above online update methods can cover large appearance changes, Occlusions etc. Even though online updating errors lead to drifting.

Hence to avoid the above mentioned problem, semi supervised AdaBoost classifier [15] is considered for tracking. It significantly avoids the drifting problem. It is also adaptive to appearance changes.in this method, the updating is performed in a semi-supervised manner.

In paper [16] visual tracking with multiple instance learning is performed.it works based on boosting. It can, therefore, lead to a more robust tracker with fewer parameter tweaks.it presents a novel online MIL algorithm for object tracking that achieves superior results with real-time performance. [17] consist of a mechanism which combining both semi-supervised learning and MIL classifier to use the previous information properly [18] presented an ensemble of linear SVM classifier with online updating. It will select major frames of the target.

3.3 Methods based on online Matric Learning

Learned classifiers are used for proper matching and tracking. In the discriminative cases, there fixed distances metrics are specified in advance. Some examples are Euclidean metrics, Bhattacharya coefficient, matusita matric a combination of them and so on.

4. COMBINED METHODS

Generative appearance models will perform well for appearance changes in subsequent frames.Eventhough it is difficult to gather a better result for the cluttered background.in discriminative appearance models, the main problem is the influence of noise or drifting. Therefore, researches make use of the simultaneous use of these two models so as to get a proper result.

Such a combined method discriminative generative model is applied in a different method. In basic methods, an Observation model built online for visual tracking and the prediction of target location is possible by using discriminative model.

5. CONCLUSION

By the proper comparative study of the appearance modeling techniques in visual tracking, it is purely clear that there is equal importance for both the discriminative and generative methods in visual tracking. The choosing of modeling will be different based on application.

6. REFERENCES

[1] Jepson, A. D., Fleet, D.J., EI-Maraghi, and TF: 'Robust online appearance models for visual tracking'. Proc, IEEE Conf Computer Vision and pattern recognition (CVPR),Kauai, HI, USA, December 2001,pp. 415-422

[2]Tu, J.I., Tao, H.: 'Online updating appearance generative mixture model for mean shift tracking'. Proc. Asian. Conf. Computer Vision (ACCV) Hyderabad, India, January 2006, pp. 694-703

[3]Zhou,S., Chellappa, R., Moghaddam, B.: 'Visual tracking and recognition, using appearance-adaptive models in particle filters', IEEE Trans.Image process. 2004, 13,(11),pp.1491-1506

[4]Ross,D., Lim,J.,Yang,M.H.: 'Adaptive probabilistic visual tracking with incremental subspace update'.Proc.European Conf.Computer Vision (ECCV),Prague, Czech Republic,May 2004,pp.470- 482

[5] Lim ,J., Ross, D.,Lin,R.S.,Yang, M.H.: 'Incrimental learning for visual tracking',in Weiss,Y.,Bottuou,L.,(Eds.), 'Advances in neural informa,tion processing systems,(MTI Press,2005),pp.793-800

[6]Lee, K., Kriegman. D: 'Online learning of probabilistic appearance manifolds for video-based recognition and tracking.' Proc IEEE Conf Computer Vision and Pattern Recognition (CVPR), San Diego,USA,June 2005, pp.852-859

[7] Xue Mei, Haibin Ling, Yi Wu, Erik Blasch, and Li Bai, "Minimum error bounded efficient l1 tracker with occlusion detection," in Proceedings of IEEE Conference on Computer Vision and pattern recognition ,2011.

[8] C. Bao, Y.Wu, H. Ling, and H. Ji, "Real time robust l1 tracker using accelerated proximal gradient approach," in IEEE Conference on Computer Vision and Pattern Recognition, 2012.

[9] Hanxi Li and Chunhua Shen, "Robust real-time visual tracking with compressed sensing," in Proceedings of IEEE International Conference on Image Processing, 2010.

[10] Xu Jia, Huchuan Lu, and Ming-Hsuan Yang, "Visual tracking via adaptive structural local sparse appearance model," in IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 1822–1829.

[11] Kaihua Zhang, Lei Zhang, and Ming-Hsuan Yang, "Real-time compressive tracking," In processdings of European Conference on Computer Vision, 2012

[12] Avidan, S.: 'Ensemble tracking'. Proc. IEEE Conf. Computer Vision and pattern recognition (CVPR), San Diego, CA, USA, June 2005, pp.494-501

[13] Leistner,C., Granber, H., Bischof, H.: 'Semi-supervised boosting using visual similarity learning'. Proc.IEEE Conf. Computer Vision and Pattern Recognition(CVPR), Anchorage,Alaska,USA,June 2008,pp.1-8

[14]Babenko,B.,Yang,M.H.,Belongie,S.: 'Visual tracking with online multiple instance learning'.Proc.IEEE Conf. CVPR,Miami, Florida, USA, June 2009,pp.983-990

[15] Amit Adam, Ehud Rivlin, and Ilan Shimshoni, "Robust fragments-based tracking using the integral histogram," in IEEE Conference on Computer Vision and Pattern Recognition, 2006, pp. 798–805.

[16] Wei Zhong, Huchuan Lu, and Ming-Hsuan Yang, "Robust object tracking via sparsity-based collaborative model," in IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 1838–1845.

[17] Zeisl,B., Leistner,C., Saffari,A., Bischof, H.: 'on-line semi-supervised multiple boosting'. Proc. IEEE Conf. CVPR , San Francisco, CA, USA, June 2010, pp.1879-1886

[18]Iian, M., Zhang,W.,Liu,F.: 'Online ensemble SVM for robust tracking'. Prof.Conf.ACCV,Tokyo, Japan,November 2007, pp.355-364

