Efficient Occlusion Handling Object Tracking System

Mr. C. Sathish1, Merwin Amala Roger.J2,
Assistant Professor, Department of CSE, Regional Centre of Anna University, Tirunelveli.
csathishme@yahoo.co.uk

PG Scholar, Department of CSE, Regional Centre of Anna University, Tirunelveli.
merwinroger@gmail.com

Abstract

Most object tracking algorithms use discriminative classifier which separates the target object from its background. This might lead to the inclusion of noisy samples when they are not sampled properly, causing tracking drift. This project work handles appearance change caused by intrinsic and extrinsic factors. A novel online object tracking algorithm is proposed which combines Principal Component Analysis (PCA) and L1 regularization. PCA is exploited with recent sparse representation schemes for learning effective appearance models. L1 regularization is introduced into the PCA reconstruction. The combination of this novel algorithm is developed to represent an object by sparse prototypes that accounts explicitly for data and noise. Occlusion and motion blur is taken into account in order to reduce tracking drift. It does not simply include the image observations for model update. Numerous experimental evaluations on challenging sequences demonstrate that the proposed tracking algorithm performs favorably well against several state-of-the-art algorithms.

Key words: Object tracking, discriminative classifier, tracking drift, Principal Component Analysis (PCA), L1 regularization, occlusion.

1. Introduction

Object tracking is one of the most crucial task in computer vision. It has its applications in many areas, especially in motion analysis. Despite of its usefulness, developing a robust tracker is a challenging problem. Because the target object is subjected to appearance change due to numerous reasons such as camera motion, pose variation, varying illumination and occlusion.

With the dynamic movement of the object, the tracker must observe the object and then track the object over each particular state of time. A search strategy must be maintained to find the likely states of the object. Developing a robust object tracker [5][10][20][21] is still a challenging problem due to difficulties that explicitly account for the appearance change of the target object. The difficulties can be intrinsic or extrinsic. The factors which correspond to intrinsic difficulties are pose variation and shape deformation. The extrinsic difficulties include some factors such as varying illumination, camera motion and occlusion.

This project work is proposed to track an object efficiently despite of the occlusion and motion blur. There are several factors that should be considered in order to develop an efficient object tracker. The most important thing to be considered is how the object is represented. The representation of an object can be based on its features such as intensity, color, texture, etc.

The next thing to be considered is whether the representation is generative or discriminative. Generative method is used to search for image regions with minimal reconstruction error. When limited data is available generative method achieves higher generalization. Whereas, discriminative methods handle the object with a classifier which separates the target object from its surrounding background. It defines a boundary and distinguishes the object from its background.

Numerous tracking algorithms have been developed within the Multiple Instance Learning (MIL) [24] framework. This project work proposes a robust generative tracking method with appearance model. It mainly handles partial occlusion and motion blur. This method does not require any complex combination of generative and discriminative trackers for handling partial occlusion [28][29].

The combination of PCA [11][22] reconstruction and L1 regularization is used for the efficient tracking of the target object. Numerous experiments and evaluations on challenging image sequences demonstrate that the proposed tracking method is efficient and effective for robust object tracking when compared to several state-of-the-art methods [1-3][6][8][23][25].

2. Related Work

Many works has been done in object tracking and more thorough reviews on this topic can be found in [30].

A. Object Tracking using Sparse Representation
Sparse representation is one of the best techniques which has its application in several fields such as computer vision and pattern recognition. Motivated by the work of [31], [21] propose an algorithm called l1 tracker. It handles the tracking problem by finding the most likely patches with sparse representation. Thus it handles partial occlusion using trivial templates. Each of the candidate image patches is sparsely represented by a set of trivial templates and thus solved using l1 minimization [21].

B. Existing System

There have been several algorithms proposed for the purpose of tracking a target object. Especially in the past decade, MIL tracking model is proposed which forms the base of all the tracking algorithms. Using this MIL tracker as a base, numerous methods have been developed with several techniques. But all the existing systems have the major problem with the tracker in the sense that the tracker must be initialized in multiple locations. And if occlusions are present, a more sophisticated motion filter should be employed. The existing systems are highly time consuming when the model complexity is high. It occasionally drifts from the target object. It also leads to poor performance in the presence of significant appearance change. It is not expected to handle appearance change due to large variation in scale and shape deformation. Also it may fail when there is similar objects in the scene. Finally all the existing trackers rely on the smoothness in appearance change and cannot deal with abrupt appearance change.

C. Motivation of this work

The main concept of this project work is to handle appearance change which is done by modeling object appearance. Both subspace online learning [4][13][19] and sparse representation are combined for this purpose. Introducing l1 regularization with PCA gains more efficiency. The major problem of occlusion can be handled by the use of trivial templates.

In this method, both the target and the trivial templates must be sparse, because the target templates are coherent in nature. The coefficients of trivial templates are used to model partial occlusions. But PCA basis vectors are not coherent in nature. Hence the coefficients of the basis vectors are not sparse but they are orthogonal. The observation of the target object can be sparsely represented by prototypes when the number of trivial templates is much larger than the number of basis vectors.

Compared with IVT tracker [5], this method handles partial occlusion efficiently and effectively. Compared with l1 tracker [21], this method handles high resolution image patches with less computational complexity. This is achieved by exploiting the subspace representation.

---

Algorithm 1 Algorithm for Computing \( z_{\text{opt}} \) and \( e_{\text{opt}} \)

Input: An observation vector \( y \), orthogonal basis vectors \( U \) and a small constant \( \lambda \).

1: Initialize \( e_0 = 0 \) and \( i = 0 \)
2: Iterate
3: Obtain \( z_{i+1} \) via \( z_{i+1} = U^T (y - e_i) \)
4: Obtain \( e_{i+1} \) via \( e_{i+1} = S_\lambda (y - Uz_{i+1}) \)
5: \( i \leftarrow i + 1 \)
6: Until convergence or termination

Output: \( z_{\text{opt}} \) and \( e_{\text{opt}} \)

---

3. Proposed System

The proposed system has a robust generative tracking algorithm with adaptive appearance model which handles partial occlusion and other challenging features. Compared with part-based models, this algorithm maintains a holistic appearance information and therefore provides a compact representation of the tracked target. By exploiting the advantage of the subspace representation, this system is able to process higher resolution image observations based on sparse representation of templates. In comparison to the subspace based tracking algorithms, this system is able to deal with heavy occlusion effectively.

This method combines PCA and l1 regularization. PCA is mainly used for tracking an object. It is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. It uses covariance matrix and finally finds the Eigen vector. L1 regularization is used for tracking an object irrespective of the obstacles. Regularization is the process of introducing some additional information in order to
prevent over-fitting. Initially the regularization is set to a constant value.

A. Models

In order to track an object efficiently, we need to consider 3 models. The first is the dynamic model. For this, an affine image wrap is applied to the target motion between two consecutive frames. The second is the observation model, which facilitates the observation of the target object from frame to frame. An image observation can be assumed to be generated from a subspace of the target object when there is no occlusion. In this case, the most likely image patch can be effectively represented by the PCA basis vector. And the coefficients related to the trivial templates tends to be zero. Also the candidate patch does not correspond to the true target location. Two parts must be considered for this purpose. One is the reconstruction error of the unoccluded proportion of the target object and the other is to penalize the labelling of any pixel that is being occluded. The third and final model is to update the observation model. This is required to handle appearance change of a target object for visual tracking [7][9][26][30]. The major problem of tracking drift occurs when the model degrades. This occurs when some imprecise samples are used for update. So we explore the trivial coefficients for occlusion detection, since the corresponding templates are used to account for noise.

B. Discussion

The result of the combination of the algorithms make our tracker robust and efficient. It handles the potential outliers, i.e., occlusion and mis-alignment more effectively. Our tracker handles partial occlusion explicitly and facilitate it to choose the well-aligned observation. Also our tracker alleviates the problem caused by inaccurate samples.
4. Experiment

The proposed algorithm is implemented in MATLAB which runs 2 frames per second in the environment of Windows 8 with 64 bit processor. Each image observation is normalized to 32 x 32 pixels for the purpose of PCA reconstruction. The regularization constant is set to 0.05.

4.1 Experiment

The proposed algorithm is implemented in MATLAB which runs 2 frames per second in the environment of Windows 8 with 64 bit processor. Each image observation is normalized to 32 x 32 pixels for the purpose of PCA reconstruction. The regularization constant is set to 0.05.

A. Qualitative Evaluation

**Heavy occlusion:** In the occlusion 1 sequence, our algorithm performs preferably well against several other methods. The proposed tracking algorithm handles occlusion using sparse representation with trivial templates. For the occlusion 2 sequence, when the object is subjected to partial occlusion or in-plane rotation, our tracker performs well. This tracker performs well in terms of position and scale even when the target is heavily occluded.

**Illumination change:** This method does not drift away when drastic illumination variation occurs. Appearance change of an object can be well approximated by a subspace at fixed pose. If the tracked target undergoes illumination variation, the appearance variation can be well modelled by a low dimension PCA subspace. For the car 4 sequence, there is a drastic lighting change. Our tracker tracks the target object efficiently even though the target object is small with low contrast and drastic illumination change in the car 11 video image set [27].

**Fast motion:** As the object undergoes abrupt motion, it is difficult to predict its location. In addition, it is rather challenging for appearance change caused by motion blur. Accurate locations of the tracked objects can be obtained by penalizing the sparsity of the error term. Our tracker efficiently tracks the target object since it properly updates the appearance model.

**Cluttered background:** There arises a challenging situation when the object undergoes change of scale and pose, as well as heavy occlusion in cluttered background. The proposed tracker adapts to scale change, in-plane rotation and occlusion. The error terms of the mis-aligned ones are larger and the corresponding representations are denser.

![Fig. 3. Tracking method](image)

![Fig. 4. Qualitative evaluation. (a) Occlusion 1. (b) DavidOutdoor. (c) Cliflbar.](image)
B. Quantitative Evaluation

The difference between the predicted and the ground truth location of the target object, as well as its average values constitute to the quantitative evaluation of the target object. The algorithm of the proposed segmentation tracking method [12] achieves lowest tracking error in almost all the sequences. The score of the target object is calculated which is the overlap rate of the target object. When the score is above 0.5, the object is regarded as being successful. A sparse based tracker known as SPCA tracker [14-18] is implemented which performs favorably against the other algorithms. The tracking result is used to update the observation model directly when the target is well tracked and the occlusion rate is very small. The occlusion rates are consequently higher when the occlusion map reflects the situation where the tracked target suffers from partial occlusion. The occlusion rates are high and the observed image patches are discarded without update if tracking returns are off the targets.

<table>
<thead>
<tr>
<th>Image sequence</th>
<th># Frames</th>
<th>Challenging factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occlusion 1 [10]</td>
<td>898</td>
<td>partial occlusion</td>
</tr>
<tr>
<td>Occlusion 2 [7]</td>
<td>819</td>
<td>partial occlusion, in-plane rotation, out-plane rotation</td>
</tr>
<tr>
<td>Caviar 1 [41]</td>
<td>382</td>
<td>partial occlusion, scale change</td>
</tr>
<tr>
<td>Caviar 2 [41]</td>
<td>500</td>
<td>partial occlusion, scale change</td>
</tr>
<tr>
<td>Car 4 [5]</td>
<td>659</td>
<td>illumination variation, scale change</td>
</tr>
<tr>
<td>Singer 1 [40]</td>
<td>321</td>
<td>illumination variation, scale change</td>
</tr>
<tr>
<td>David Indoor [5]</td>
<td>462</td>
<td>illumination variation, scale change, out-plane rotation</td>
</tr>
<tr>
<td>Car 11 [5]</td>
<td>393</td>
<td>illumination variation, scale change, background clutter</td>
</tr>
<tr>
<td>Deer [40]</td>
<td>71</td>
<td>abrupt motion, background clutter</td>
</tr>
<tr>
<td>Jumping [25]</td>
<td>313</td>
<td>abrupt motion</td>
</tr>
<tr>
<td>Lemming [19]</td>
<td>1336</td>
<td>out-plane rotation, scale change occlusion, background clutter</td>
</tr>
<tr>
<td>Cliffbar [7]</td>
<td>471</td>
<td>in-plane rotation, scale change, background clutter, abrupt motion</td>
</tr>
</tbody>
</table>

Tab I. Evaluated Image Sequences

5. Conclusion

A robust tracking algorithm with sparse prototype representation is proposed in this project work. It resolves occlusion and object tracking is done by exploiting the strength of the sparse representation. It tracks an object efficiently irrespective of its size, position and obstacles among multiple objects. This project work mainly focuses and explicitly handles partial occlusion and motion blur. It is done for the purpose of appearance update. The object tracking is achieved by exploiting the strength of subspace model and sparse representation. Since the proposed algorithm involves handling l1 regularization, more efficient algorithms for real time applications can be explored. Experiments on several image sequences demonstrate that this tracking method performs favorably well against several state-of-the-art methods. This representation scheme can be extended for other vision problems including object recognition. Further this project work can also be extended for integrating multiple visual cues for the better description of objects in different scenarios. It can also be used to utilize prior knowledge with online learning for more effective object tracking.

References


