3rd International Conference on Emerging Trends in Engineering, Technology, Science and Management
(IETE) Institution of Electronics and Telecommunication Engineers, New Delhi, India
11th June 2017, www.conferenceworld.in
(ICETETSM-17)
ISBN: 978-93-86171-48-1

VISUAL OBJECT TRACKING AND DETECTION IN COMPUTER VISION

Mr. S. S. Athatey[1], Prof. Y. M. Kurwade[2], Dr. V. M. Thakare[3]

SGBAU, Amravati, (India)

ABSTRACT

Visual tracking is one of the most important topics in computer vision due to its wide potential applications, such as motion analysis, video surveillance, and human–computer interaction. This paper proposed a method for visual tracking which is based on local sparse descriptors and DML tracker. To detect target from the background positive and negative training samples are extracted from the target region and the background regions respectively. Then these samples are input to DML tracker, with the help of feed forward neural network and distance metric DML tracker can find true match.

Index Terms – DML tracker, local descriptor, DSS map, visual tracking

I. INTRODUCTION

Visual tracking is one of the most important topics in computer vision due to its wide potential applications, such as motion analysis, video surveillance, and human– computer interaction.Visual object tracking still remains a challenging problem in computer vision due to large appearance variations in visual objects such as varying deformations, illuminations, out-of-plane rotations, occlusions, and cluttered backgrounds [1]. Generally, the distribution of objects is usually in a nonlinear manifold due to various variations such as deformations, illuminations, and occlusions, and hence, it is desirable to employ nonlinear discriminative methods to exploit such information [3].Generative methods focus on searching for the regions that are the most similar to the tracked targets, including template-based, subspace-based, and sparse representation-based models. While discriminative methods cast tracking as a classification problem that distinguishes the tracked targets from the surrounding backgrounds [2],[3].Sparse coding based appearance models have been successfully applied to visual tracking [3].The main problem with current pooling methods for visual tracking lies in the fact that final features considering the sparse codes of all patches are usually sensitive to occlusion and other interfering factors. When the target undergoes a long-time partial occlusion or heavy deformation, performance of these methods would degrade greatly.With DSS map, candidates are evaluated in both directions: not only how similar it is to the target object but also how different it is from the background [4].A nonparametric data-driven local metric adjustment methodminimizes the deviation of the empirical misclassification probability to obtain the optimal metric such that the asymptotic error as if using an infinite set of training samples can be approximated [5].

This paper, first take overview of five existing methods for visual tracking and object detection and proposed a methodology to detect and track the object from image. The proposed method combine the concept of local sparse descriptors with DML tracker. The sparse descriptors are used to find the target object from scene. Also

structural local descriptors with local patches that are not occluded can achieve good discrimination. Descriptors of positive and negativetraining samples are extracted from the target region and thebackground regions respectively.Then it is input to DML tracker. DML tracker then track object and find out the true match.

II.BACKGROUND

A deep metric learning (DML) approach for robust visual tracking under the particle filter framework. The proposed DML tracker can explicitly learn several hierarchical nonlinear transformations to map data points into another subspace via feed-forward neural network architecture so that these nonlinear transformations are explicitly solved by maximizing the interclass variations of negative pairs and minimizing the intra-class variations of positive pairs simultaneously [1].

A novel linear regression method, least soft-threshold squares (LSS), which assumes that the error vectors follow the independent identically distributed Gaussian–Laplacian distribution. Second, an efficient iteration method to solve the LSS problem and propose an LSS distances to measure the dissimilarity between the observation vector and a dictionary of basis vectors [2].

The structural local sparse descriptor to represent the target region. When the target suffers from partial occlusion, our structural local descriptors with local patches that are not occluded can achieve good discrimination. Moreover, an occlusion-aware template update scheme can handle appearance changes during tracking especially caused by occlusion. For background clutter, the proposed tracker considers the background information to train our classifiers, which ensures the good separating capacity of distinguishing the target from the cluttered background [3].

A reversed multitask sparse tracking framework which projects the templates matrix (both positive and negative templates) into the candidates space. By selecting and weighting the discriminative sparse coefficients, the DSS map and pooling method lead to the best candidate [4].

A nonparametric data-driven local metric adjustment method. It finds a spatially adaptive metric that exhibits different properties at different locations in the feature space, due to the differences of the data distribution in a local neighborhood. It minimizes the deviation of the empirical misclassification probability to obtain the optimal metric [5].

The rest of the paper is organized as follows. In this paper, **Section II** gives us background details, **Section III** provides work which is done previously, **Section IV** gives idea about existing technology, in **Section V** analysis and discussion about techniques is carried out, proposed methodology is explained in **Section VI**, Possible outcomes and Result is described in **Section VII**, **Section VIII** concludes the paper.Finally, **Section IX** described future scope of the paper.

III. PREVIOUS WORK DONE

Junlin Hu *et.al.*(2016)[1] proposed a deep metric learning (DML) approach for robust visual tracking under the particle filter framework.The DML tracker can explicitly learn several hierarchical nonlinear transformations to map data points into another subspace via feed-forward neural network architecture so that these nonlinear transformations are explicitly solved by maximizing the interclass variations of negative pairs and minimizing

the intra-class variations of positive pairs simultaneously. It overcomes both the nonlinearity and scalability problems of conventional metric learning methods and kernel based method.

Dong Wang *et.al.* (2016)[2] present a robust tracking algorithm based on linear regression. Also introduce a linear regression method and least soft-threshold squares (LSS) method is introduced. The observation likelihood of each candidate is computed based on the LSS distance and is improved by introducing negative templates.

Bo Ma *et.al.* (2014)[3] proposes the structural local sparse descriptor to represent the target region. A target represented by using the collection of local sparse descriptors, where each descriptor represents partial appearance of the target.For background clutter, the proposed tracker considers the background information to train our classifiers, which ensures the good separating capacity of distinguishing the target from the cluttered background.

BohanZhuanget.al.(2014)[4]proposed a reversed multitask sparse tracking frameworkwhich projects the templates matrix (both positive and negative templates) into the candidates space. By selecting and weighting the discriminative sparse coefficients, the DSS map and pooling method lead to the best candidate.With this DSS map, candidates are evaluated in both directions: not only how similar it is to the target object but also how different it is from the background.

Nan Jiang *et.al.*(2014)[5] proposes a nonparametric data-driven local metric adjustment method. It finds a spatially adaptive metric that exhibits different properties at different locations in the feature space, due to the differences of the data distribution in a local neighborhood.It minimizes the deviation of the empirical misclassification probability to obtain the optimal metric such that the asymptotic error as if using an infinite set of training samples can be approximated.

IV. EXISTING METHODOLOGIES

A. Robust tracking algorithm :- It is based on linear regression. A novel linear regression method, least soft-threshold squares (LSS), which assumes that the error vectors follow the independent identically distributed (i.i.d.) Gaussian–Laplacian distribution. Second, an efficient iteration method to solve the LSS problem and propose an LSS distances to measure the dissimilarity between the observation vector and a dictionary of basis vectors. Third, design a robust tracker using the LSS method, where the dictionary consists of principal component analysis (PCA) basis vectors. The observation likelihood of each candidate is computed based on the LSS distance and is improved by introducing negative templates. Furthermore, update the tracker using an effective update scheme and speed up the tracker using an particle selection mechanism.The proposed LSS regression is equivalent to robust regression with the Huber loss function

$$\mathbf{x} = \arg\min_{\mathbf{x}} \sum_{i=1}^{d} f(e_i), \quad e_i = y_i - \mathbf{r}_i \cdot \mathbf{x}$$

B. Structural local sparse descriptors :- In this method a target representation using the collection of local sparse descriptors, where each descriptor represents partial appearance of the target. When the target suffers from partial occlusion, our structural local descriptors with local patches that are not occluded can achieve good discrimination. Moreover, an occlusion-aware template update scheme can handle appearance changes during

tracking especially caused by occlusion. For background clutter, the proposed tracker considers the background information to train our classifiers, which ensures the good separating capacity of distinguishing the target from the cluttered background.

Input: sequence $o_1, \cdots, o_M$; object state s_1 at the first frame; the number of templates n; update frequency u;
Output: tracking results s_t, $t = 2, \cdots, M$.
Initialization:
template set T and positive training set N_p from the first n frames; negative training set N_q from the nth frame; reconstructed target set $\Phi = \emptyset$.
extract local descriptors for $[N_p, N_q]$.
construct the appearance model using local descriptors by the strong classifier $H(x)$ with (3).
while $t = n+1, \cdots, M$ **do**
  generate candidate targets X.
  extract local descriptors for X.
  calculate classification score and weight for X.
  select tracking result s_t using (16).
  draw p positive samples $\Rightarrow N_p$.
  reconstruct s_t using (12) and (13), and then $\Phi = [\Phi, s_t]$.
  if size$(\Phi) == u$ **then**
    update U and μ using Φ, then $\Phi = \emptyset$.
    reconstruct T^* using (14) and update T.
    draw q negative samples $\Rightarrow N_q$.
    extract local descriptors for $[N_p, N_q]$.
    retrain $H(x)$.
  end if
end while

Algorithm1 :- Local Sparse descriptor algorithm

C. Reversed multitask sparse tracking framework :- It projects the templates matrix (both positive and negative templates) into the candidates space. By selecting and weighting the discriminative sparse coefficients, the DSS map and pooling method lead to the best candidate. A discriminative sparse similarity map (DSS map) based upon that similarity relationship. The discriminative information containing in this map comes from a large template set composed by multiple positive target templates and hundreds of negative templates. With this DSS map, candidates are evaluated in both directions: not only how similar it is to the target object but also how different it is from the background.

D. Nonparametric data-driven local metric adjustment method :- In this method a data-driven and local approach to metric adjustment in the context of visual target tracking. A global metric is such that is stationary (or the same) at all locations in the feature space, but local metric is spatially-adaptive, i.e., it varies at different locations depending on the data distribution in their local regions. It finds a spatially adaptive metric that exhibits different properties at different locations in the feature space, due to the differences of the data distribution in a local neighborhood. It minimizes the deviation of the empirical misclassification probability to obtain the optimal metric such that the asymptotic error as if using an infinite set of training samples can be approximated.

(ICETETSM-17)
ISBN: 978-93-86171-48-1

E. Deep metric learning (DML) :- It is based on the particle filter framework. The DML tracker can explicitly learn several hierarchical nonlinear transformations to map data points into another subspace via feed-forward neural network architecture so that these nonlinear transformations are explicitly solved by maximizing the interclass variations of negative pairs and minimizing the intra-class variations of positive pairs simultaneously. It overcomes both the nonlinearity and scalability problems of conventional metric learning methods and kernel based method.

Input: Training data: $\mathbf{Y} = \{(\mathbf{y}_i, \mathbf{y}_j, \ell_{ij})\}$; Layers of network: $\mathcal{K}+1$; Trade-off parameter: α; Regularization parameter: β; The learning rate: ρ; Total iterative number: T; Convergence error ε.
// Optimization procedure of DML
$\mathcal{O}_0 \leftarrow 0$;
Initializing $\{\mathbf{W}^{(k)}, \mathbf{b}^{(k)}\}_{k=1}^{\mathcal{K}}$ according to (18);
for $t = 1, 2, \cdots, T$ **do**
  for $k = 1, 2, \cdots, \mathcal{K}$ **do**
    Computing hierarchical representation $\mathbf{y}_i^{(k)}$ of each sample by using forward propagation;
  end
  // Back propagation
  for $k = \mathcal{K}, \mathcal{K}-1, \cdots, 1$ **do**
    Obtaining $\partial\mathcal{O}/\partial\mathbf{W}^{(k)}$ and $\partial\mathcal{O}/\partial\mathbf{b}^{(k)}$ in line with (9) and (10), respectively;
  end
  // Updating parameters
  for $k = 1, 2, \cdots, \mathcal{K}$ **do**
    $\mathbf{W}^{(k)} \leftarrow \mathbf{W}^{(k)} - \rho\partial\mathcal{O}/\partial\mathbf{W}^{(k)}$;
    $\mathbf{b}^{(k)} \leftarrow \mathbf{b}^{(k)} - \rho\partial\mathcal{O}/\partial\mathbf{b}^{(k)}$;
  end
  Calculating objective $\mathcal{O}_t$ using (8);
  If $|\mathcal{O}_t - \mathcal{O}_{t-1}| < \varepsilon$, go to **Output**;
end
Output: Weights and biases: $\{\mathbf{W}^{(k)}, \mathbf{b}^{(k)}\}_{k=1}^{\mathcal{K}}$.

Algorithm2 :- DML Tracker

V. ANALYSIS AND DISCUSSION

This section present some observation about the above existing visual tracking method.The table shows the comparison between five existing methods and also shows advantages and disadvantages of five methods.

Object Detection and Tracking Technique	Advantages	Disadvantages
Robust tracking algorithm	The propose a robust tracker based on the linear representation model, which is able to handle the outliers and therefore effectively provides an accurate match. The propose algorithm gives more accuracy than compared tracking algorithm. The propose algorithm have better performance than other.	It is not useful in offline tracking. When frame rate is more than 10frames/s then algorithm outperform.
Structural local	The tracker achieves the best performance among all	The target undergoes a long-time partial occlusion

sparse descriptors	compared trackers. Algorithm performs 7.0% better than Struck, 7.6% better than SCM. The method gene- rates the structural local descriptors for the target with good discrimination.	or heavy deformation, performance of these methods would degrade greatly.
Reversed multitask sparse tracking framework	The Laplacian constraint serves to increase the stability of the proposed algorithm. The algorithm is better than OWN and OWL algorithm in challenging scene.	Without negative templates or the Laplacian constraint, the robustness of proposed tracker indeed decreases to some extent.
Nonparametric data-driven local metric adjustment method	The method have the better performance over other method's. It take the data local distribution into consideration this new local metric learning method into target tracking leads to efficient and robust tracking performance.	In the proposed metric adjustment method, one critical issue is the determination of right scale for the local region, because different volumes of the local regions cover different sets of training samples, and thus different local data distributions.
Deep metric learning (DML)	DML tracker first learns a set of hierarchical nonlinear transform- ations in the feed forward neural network. The performance of proposed method is slightly better than other compared methods.	Fast motion and motion blur are two very challenging factors for the DML tracker. The tracking performance of DML gradually reduces when the sample dimensionality is too low or too high.

TABLE 1: Comparisons between different Object Detection and tracking techniques

VI. PROPOSED METHODOLOGY

This section, provides the proposed methodology which have used a concept of deep metric learning and sparse coding method. Sparse coding based appearance models have been successfully applied to visual tracking. The proposeda target representation using the collection of local sparse descriptors, where each descriptor represents partial appearance of the target. Descriptors of positive and negative training samples are extracted from the target region and the background regions respectively. For tracking purpose use the DML tracker. As DML tracker can explicitly learn several hierarchical nonlinear transformations to map data points into another subspace via feed-forward neural network architecture so that these nonlinear transformations are explicitly solved by maximizing the interclass variations of negative pairs and minimizing the intra-class variations of positive pairs simultaneously.

Algorithm:- Combination of sparse descriptor with DML tracker for object tracking and detection.

Step1: Extract structural local sparse descriptors from a fraction of all patches by performing a pooling operation.

(ICETETSM-17)
ISBN: 978-93-86171-48-1

Step 2: To find the target from background extract descriptors of positive and negative training samples from target region and background regions.

Step 3: Apply the DML tracker approach to find or identify a true target.

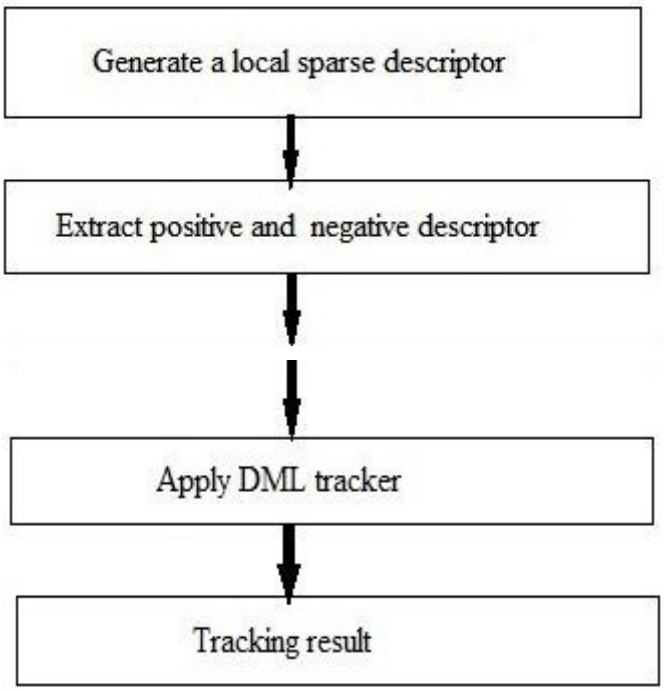

Fig 1:- Process flow diagram

VII. POSSIBLE OUTCOMES AND RESULTS

When the target suffers from partial occlusion, structural local descriptors with local patches that are not occluded can achieve good discrimination. Also, an occlusion-aware template update scheme can handle appearance changes during tracking especially caused by occlusion. For tracking purpose use the DML tracking approach, it overcomes the nonlinearity issue of conventional metric learning methods and scalability issue of kernel based method by using feed forward neural network architecture.

VIII. CONCLUSION

This paper, proposed a method which includes sparse coding and DML tracker in order to track accurate target from scene. Combination of sparse coding and DML tracker improves visual tracking by overcoming the issues like nonlinearity and scalability. Also sparse descriptor considers the background information to train our classifiers, which ensures the good separating capacity of distinguishing the target from the cluttered background.

IX. FUTURE SCOPE

From the observation found that the fast motion and motion blur are two very challenging factors for the DML tracker. Also tracking performance of DML gradually reduces when the sample dimensionality is too low or too high. Future study tries toovercome this issue to get high performance.

REFERENCES

[1] Junlin Hu, Jiwen Lu and Yap-Peng Tan, Senior Member, IEEE, “Deep Metric Learning for Visual Tracking”, *IEEE Transaction on Circuits and Systems for video technology,* Vol. 26, No. 11, PP. 2056-2068, November 2016.

[2] Dong Wang, Huchuan Lu and Ming-Hsuan Yang, Senior Member, IEEE, "Robust Visual Tracking via Least Soft-Threshold Squares", *IEEE Transaction on Circuits and systems for video technology,* Vol. 26, No. 9, PP. 1709-1721, September 2016.

[3] Bo Ma, JianbingShen, Yangbiao Liu, Hongwei Hu, Ling Shao and Xuelong Li, "Visual Tracking Using Strong Classifier and Structural Local Sparse Descriptors", *IEEE Transaction on Multimedia,* Vol. 23, Issue no. 4, Page. 1872-1881, April 2014.

[4] BohanZhuang, Huchuan Lu, Senior Member, IEEE, Ziyang Xiao, and Dong Wang, "Visual Tracking via Discriminative Sparse Similarity Map", *IEEE Transaction on Image Processing,* Vol. 23, No. 4, PP. 1872-1881, April 2014.

[5] Nan Jiang and Wenyu Liu, Member, IEEE, "Data-Driven Spatially-Adaptive Metric Adjustment for Visual Tracking", *IEEE Transaction on Image Processing,* Vol. 23, No. 4, PP. 1556-1568, April 2014.