Visual Tracking via Sparse Coding and Spectral Residual

Wei Li
College of Jin Cheng
Nanjing University of Aeronautics and Astronautics
Nanjing, China
e-mail: spvivi@hotmail.com

Meng Ding
College of Civil Aviation
Nanjing University of Aeronautics and Astronautics
Nanjing, China
e-mail: nuaa_dm@hotmail.com

***Abstract*— This paper proposed a tracking algorithm based on sparse coding and spectral residual saliency under the framework of particle filtering. The proposed algorithm can be divided into three parts. Firstly, spectral residual is used to calculate a saliency map of the current frame and then compute the saliency score of each particle. Secondly, several particles are eliminated directly based on the differences between the saliency scores of the particles in the current frame and the target score in the prior frame. Thirdly, ScSPM is used to compute the observation vector for the rest particles and the tracking task is finished in the framework of particle filtering. Both quantitative and qualitative experimental results demonstrate that the proposed algorithm performs favorably against the nine state-of-the-art trackers on ten challenging test sequences.**

Keywords—visual tracking, sparse coding, spectral residual saliency.

I. INTRODUCTION

For a successful and effective visual tracking algorithm, building an appearance model is of great importance. Currently, most of state-of-the-art tracking algorithms are driven by two appearance models [1]: discriminative model-based tracker and generative model-based tracker. The essence of discriminative model-based tracker is a binary classification problem and it usually contains two steps [2-10]: feature representation and classifier training. The classifier must be updated by online training algorithms, such as online boosting (OAB) [7], online supervised boosting (SemiB) [8], online multiple instance learning (MIL) [5], online weighted multiple instance learning (WMIL) [8]. Online training is formulated as a process of semi-supervised learning and may lead to the drifting problem. The generative model-based trackers typically learn appearance models to represent the tracking target and then formulate the visual tracking as an optimization problem which searches for the image patch with maximum likelihood or minimal error in the subsequent frames [11-15], such as adaptive structural local sparse appearance model (ASLSA) [11], L1 tracker using accelerated proximal gradient approach (L1APG) [12], distribution fields for tracking [13]. In most cases, the generative model-based trackers are implemented under the filtering framework, such as Kalman filter (KF) and its variants [16], or particle filter (PF) [17, 18]. Compared with KF and its variants, PF doesn't need to assume a Gaussian distribution for posterior [17].

It is commonly believed that a robust and effective appearance model of the tracking target plays an important role in generative model-based trackers. Generally, the features used in appearance models can be divided into three levels: primary, intermediate and advanced features. In recent years, with the development of sparse coding, advanced features obtained by unsupervised learning through further and deep extraction from intermediate features become extremely popular in computer vision. For example, Yang et al proposed a linear spatial pyramid matching kernel based on Scale-invariant feature transform (SIFT) sparse codes called ScSPM [19] (Sparse coding Spatial Pyramid Matching). The proposed algorithm uses ScSPM to obtain the observation vector to construct an appearance model of the tracking target. Sparse coding has attracted a great of attention in the field of visual tracking. These trackers show state-of-the-art performance as reported in previous publications [20]. Recently, several trackers combining PF and sparse coding have been presented [21]. These methods can solve the problem of target offset in traditional tracking algorithms.

In order to improve real-time performance, the proposed algorithm first uses an approach of saliency analysis with low computation cost to obtain the saliency scores of all candidate particles and then selects some of these candidate particles to extract observation vector using ScSPM according to the score of each candidate particle.

The rest of this paper is organized as follows: section 2 presents the details of our tracking algorithm using ScSPM and Spectral Residual saliency. In section 3, the experimental results and related comparisons on challenging image sequences are demonstrated. Finally, conclusion and future work are summarized in section 4.

II. PROPOSED TRACKING ALGORITHM

This section introduces the details of the proposed tracking algorithm. The framework of the proposed algorithm is shown in Fig.1.

A. *observation vector extraction using ScSPM*

Suppose an intermediate feature vector $\mathbf{y} \chi \mathbf{R}^{n}$, and an over-complete dictionary $\mathbf{D} \chi \mathbf{R}^{n \times m}$, m>>n. The dictionary used in the proposed algorithm is learned from some nature image datasets which contain similar structures to the tracking

978-1-5386-1978-0/17/$31.00 ©2017 IEEE