Neurocomputing 152 (2015) 345–358

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Correlation-based self-correcting tracking

Tewodros A. Biresaw [a,b,*], Andrea Cavallaro [b], Carlo S. Regazzoni [a]

[a] DITEN, University of Genova, Italy
[b] Centre for Intelligent Sensing, Queen Mary University of London, UK

ARTICLE INFO

Article history:
Received 30 August 2013
Received in revised form
30 July 2014
Accepted 26 October 2014
Communicated by Paolo Remagnino
Available online 4 November 2014

Keywords:
Object tracking
Data association
Online performance measure
Correction
Trajectory correlation

ABSTRACT

We present a framework for improving probabilistic tracking of an extended object with a set of model points. The framework combines the tracker with an on-line performance measure and a correction technique. We correlate model point trajectories to improve on-line the accuracy of a failed or an uncertain tracker. A model point tracker gets assistance from neighboring trackers whenever a degradation in its performance is detected using the on-line performance measure. The correction of the model point state is based on correlation information from the state of other trackers. Partial Least Square (PLS) regression is used to model the correlation of point tracker states from short windowed trajectories adaptively. Experimental results on data obtained from optical motion capture systems show the improvement in tracking performance of the proposed framework compared to the baseline tracker and other state-of-the-art trackers.

1. Introduction

Tracking plays a fundamental role in surveillance [1], computer vision [2], human–computer interaction [3] and medical image processing [4]. A wide variety of tracking techniques have been proposed such as Mean-Shift tracker [5], Kalman filter tracker [6] and KLT tracker [7] (surveys on tracking are available in [8–10]). The target state representation and tracking challenges vary from one application domain to another. For example, an extended object can be represented by a set of points estimated from multiple independent measurements [11,8]. The movements of the points enable us to analyze the overall shape evolution and sub-part dynamics, such as the movements of hands and legs relative to other body parts of a person [12,13]. Local appearance is modeled using pre-selected points on the object, which we refer to as *model points*, such as markers in a motion capture system [14], or features extracted from images using Scale-Invariant Feature Transform (SIFT) [15].

Tracking model points is achieved by estimating the state of each point individually, which generates a challenge for data association [11]. Moreover, performance degradation or failures in tracking can be generated by the challenges related to data association, missed and false detections, illumination changes and occlusions [16]. A tracking failure at any instant of time can generate a long-term failure due to the use of first-order Markov processes and model updates [17,18]. Except [19–21], most trackers do not explicitly detect failures and correct them. A performance measure to detect tracking failures and a tracking correction step are desirable to obtain robust tracking [22–24]. Since comparing the tracker's output to the ground-truth data is not applicable for real-time systems [25,26], there is the need for an efficient and robust framework for online track verification and correction [19,27].

We propose a Track-Evaluate-Correct (TEC) framework based on Bayesian filtering of model points. Tracking is obtained using a baseline tracker. Evaluation and correction judge the track quality and apply appropriate changes to the baseline tracker for improving its performance. As a novelty, we make model point trackers to assist each other based on their evaluation and a correlation model in the TEC framework. In particular, we propose a quality measure criterion for evaluation of each model point track to produce a decision for correction. Correction of low-quality model point tracker involves an estimation of a probable true state and a re-initialization of the tracker using the correlation model with other point trackers. The correlation between point trackers is modeled from observed trajectory histories adaptively based on the result of the quality measure. Unlike our previous work on performance evaluation using time-reversed Markov chain [28] and trajectory correlation [29], in this work we take a binary decision on the trackers by examining their states, and we use an online modeling and a correlated trajectories selection criterion for effectively

* Corresponding author at: Centre for Intelligent Sensing, Queen Mary University of London, UK.
E-mail addresses: tewodros.biresaw@eecs.qmul.ac.uk (T.A. Biresaw), andrea.cavallaro@eecs.qmul.ac.uk (A. Cavallaro), carlo.regazzoni@unige.it (C.S. Regazzoni).
URLS: http://www.eecs.qmul.ac.uk/~andrea (A. Cavallaro), http://www.isip40.it (C.S. Regazzoni).

http://dx.doi.org/10.1016/j.neucom.2014.10.057
0925-2312/

recovering low-quality trackers. The proposed TEC framework is shown in Fig. 1.

The paper is organized as follows: Section 2 formulates the problem. Related works on evaluation and correction for tracking are reviewed in Section 3. Section 4 discusses the algorithm for tracking model points. The proposed performance measure and correction technique are presented in Sections 5 and 6, respectively. Section 7 discusses the experiments and Section 8 concludes the paper.

2. Problem formulation

Tracking involves estimating the states $\mathbf{X}=\{\mathbf{X}_t\}_{t=1}^{\tau}$ of the target over time from a set of available measurements $\mathbf{Z}=\{\mathbf{Z}_t\}_{t=1}^{\tau}$. τ is the trajectory duration, while $\mathbf{X}_t$ and $\mathbf{Z}_t$ are the estimated state and the measurement, respectively, at time t. Let us define $T(\cdot)$ to represent a tracker that estimates $\mathbf{X}_t$ as

$$\mathbf{X}_t = T(\mathbf{Z}_t, \zeta_t), \tag{1}$$

where ζ_t is other inputs to the tracker such as the previous state $\mathbf{X}_{t-1}$ for Bayesian trackers (Fig. 1). For an extended object, the state $\mathbf{X}_t$ consists of the state (and identity) of each model point

$$\mathbf{X}_t = \left\{ \mathbf{x}_t^l : 1 \le l \le N_t, l \in \mathbb{N}^+ \right\}, \tag{2}$$

where $\mathbf{x}_t^l$ is the state of model point l and N_t is the number of estimated model points. For implementations with initiation and termination of model points, N_t is variable over time. Similarly, at each time the sensor or feature extractor produces M_t point measurements

$$\mathbf{Z}_t = \left\{ \mathbf{z}_t^{\hat{l}} : 1 \le \hat{l} \le M_t, \hat{l} \in \mathbb{N}^+ \right\}, \tag{3}$$

where $\mathbf{z}_t^{\hat{l}}$ is the $\hat{l}^{th}$ model point measurement. The measurements $\mathbf{Z}_t$ are unlabeled, and are affected by potential misdetections and clutter.

Individual allocated trackers for each model point are *local trackers* T^l. The state vector $\mathbf{x}_t^l$ depends on the type of motion model used in the tracking method. A typical state vector for a D-dimensional tracking problem contains the position and velocity components of the model point as $\mathbf{x}_t^l = [x_{t,1}, \dot{x}_{t,1}, \ldots x_{t,D}, \dot{x}_{t,D}]^{\mathsf{T}}$.

The quality of the tracking result depends on how close the estimated state is to the actual (true) state. A performance measure on the tracker T^l and estimated states $\mathbf{x}_t^l$ is represented as

$$p_t^l = \Phi(\mathbf{x}_t^l, T^l, I_p), \quad 1 \le l \le N_t, \tag{4}$$

where $\Phi(\cdot)$ is the operation made to obtain a set of predefined classes or numerical values p_t^l for the track quality or the tracker performance measure. I_p is any other information, other than the current estimated state of the tracker, such as pre-defined threshold values and reference data, which are used by the performance measure.

When a tracker generates a low-quality output, a correction mechanism can be employed. The correction is applied to the tracker and the track based on a decision from the result of the quality measure $\mathbf{p}=\{p_t^l\}_{l=1}^{N_t}$. For the identified low-quality tracks, let us define their performance value as $p_t^l=1$. The correction step aims to modify the tracker T^l and to improve the accuracy of the estimated states $\mathbf{x}_t^l$ as

$$\hat{T}^l, \hat{\mathbf{x}}_t^l = \begin{cases} \Theta(\mathbf{x}_t^l, T^l, p_t^l, I_c) & \text{if } p_t^l = 1, \\ T, \mathbf{x}_t^l & \text{otherwise} \end{cases} \tag{5}$$

where $\Theta(\cdot)$ is the transformation made to obtain the corrected tracker $\hat{T}$ and the improved states $\hat{\mathbf{X}}_t = \left\{ \hat{\mathbf{x}}_t^l \right\}_{l=1}^{N_t}$. I_c, which is similar to I_p, represents valid information such as a trajectory output and an online learned appearance model to assist the correction technique. Determining $\Phi(\cdot)$ and $\Theta(\cdot)$ for the tracker, together with the methods to obtain and use the side information, I_p and I_c, plays an important role in the implementation of the overall TEC framework.

3. Prior work

The idea of track quality measurement and correction is used to improve the performance of baseline trackers. Various types of performance measures and correction techniques, i.e. $\Phi(\cdot)$ and $\Theta(\cdot)$, have been proposed for different trackers, and are discussed in this section. Table 1 summarizes the characteristics of trackers that use performance measures and/or correction techniques.

3.1. Performance measures

Quantifying the quality of a tracking result generally involves comparing one or more output variables of a tracker with reference data [25]. For an offline case, manually collected ground-truth data are used as a reference [25]. For online performance measures, standalone empirical methods judge the output of the tracker [26]. The characteristics of the output considered for evaluation include trajectory properties [19,37,39], objects color differences and boundary contrasts with background [33,40,41], observation likelihood [25,32,41] and innovation errors or covariances of the states [25,31]. These characteristics are compared with thresholds and predefined target properties. Trajectory properties such as smoothness, length, change of direction, similarity with a predefined model and similarity with a reverse tracking result are considered for trajectory-based evaluation [31,39,42]. Histogram differences between the track output and the prior known reference of the target or temporal differences between track outputs are used as color-oriented performance measures.

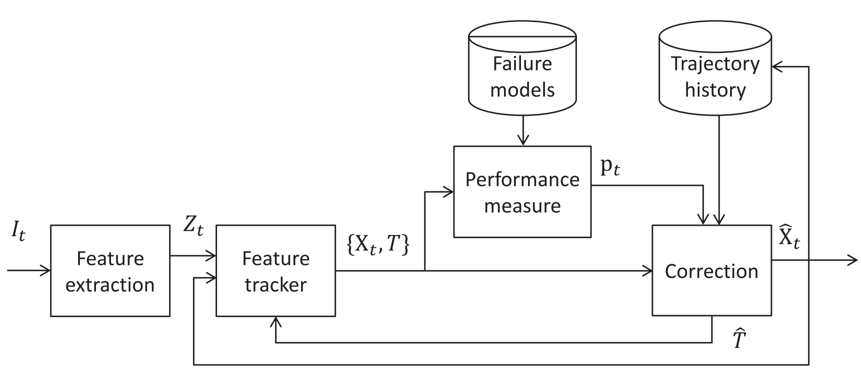

Fig. 1. Block diagram of the proposed Track-Evaluate-Correct (TEC) framework. From the image I_t, the feature extractor produces a set of measurement $\mathbf{Z}_t$ and the baseline tracker estimates the set of tracks $\mathbf{X}_t$. A performance measure on the state $\mathbf{X}_t$ and tracker T gives a set of binary decisions $\mathbf{p}_t$ that expresses the track quality. Based on $\mathbf{p}_t$ a correction step produces a more accurate estimate $\hat{\mathbf{X}}_t$ and a corrected tracker $\hat{T}$. Failure models and the knowledge of the past trajectory information allow evaluation and correction in the framework.