Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 201 (2017) 280–286

Procedia Engineering

www.elsevier.com/locate/procedia

3rd International Conference "Information Technology and Nanotechnology, ITNT-2017, 25-27 April 2017, Samara, Russia

Target tracking with composite linear filters on noisy scenes

Vitaly Kober[a,b,]*, Vladislav Kuznetsov[b]

[a]*Department of Computer Science, CICESE, Carretera Ensenada-Tijuana 3918, Zona Playitas, Ensenada, B.C. 22860, Mexico*
[b]*Chelyabinsk State University, 129 Br. Kashirinikh, Chelyabinsk, 454001, Russia*

Abstract

A tracking system using a bank of adaptive linear filters is proposed. Tracking is carried out by means of multiple target detections. The linear filters are designed from multiple views of a target using synthetic discriminant functions. For each view an optimum filter is derived from noisy reference image and disjoint background model. An iterative algorithm is used to improve the performance of the synthesized filters. The number of filters in the bank can be controlled to guarantee a prescribed tracking accuracy. Computer simulation results show that the proposed algorithm is able to precisely track a target.

Keywords: Object tracking; correlation filters; composite filters

1. Introduction

Object tracking systems are used for applications such as video surveillance, motion based recognition, and vehicle navigation [1]. Tracking requires processing large amounts of data. Two approaches can be taken: to reduce the amount of information to be processed and to carry out the processing faster. In the former approach, features are usually computed. A feature extractor ideally outputs a small number of features. Matching these features across frames yields the displacement information. When the camera rate is high, preprocessing might be done by

* Corresponding author.
E-mail address: vkober@hotmail.com

10.1016/j.proeng.2017.09.625

subtracting the background of a given frame from the next one, so that the only information is left in the frame in the area where movement took place. On the other hand, tracking can be done on the original image without extracting features, by using appearance-based models [1-4]. This approach requires that either the appearance of a target does not change much along the frame sequence or possible views of the target should be known a priori. This results in a high computational cost. Optical processors were used extensively for object detection [5]. They perform a fast detection by exploiting the parallelism inherent in optical systems. In the case when tracking requires only knowing the position of the object at a given time, the problem may be solved by using a consecutive detection approach [6]. The use of optical correlators for fast object detection allows real-time tracking applications. The approach of tracking by successive detection has several advantages. Correlation filters can be designed to analytically minimize the probability detection errors, thus detection in each scene is optimum with respect to a detection criterion. Additionally, correlation filters can be designed to minimize errors in the estimation of the target location. Furthermore, since the target is being detected in each scene, there is no problem with situations when the target is being temporarily occluded in the scene because it can be correctly detected upon reentering the scene.

Composite correlation filters were proposed for taking into account multiple views of a target in a single correlation operation [7]. An optimum correlation filter can be designed for each view of the target, and filters for multiple views can be combined into a single composite filter [6-11]. If information about a background where detection will be carried out is available, the discrimination capability (DC) of composite filters can be improved using an adaptive approach [12]. It has been shown that the performance of composing filters degrades with an increasing number of views. This problem can be solved by using a bank of composite filters when each of them is designed with a subset of known views of the target. Filters in the bank are then applied in rapid successive 8correlations, and the maximum value over all correlation output planes is chosen as the estimation of the location of the target. The number of filters in the bank is chosen to ensure a required accuracy. In other words, the parameter space of possible distortions is divided in such a way to always get an error of the position estimation less than a prescribed value with the minimum number of correlations.

The paper is organized as follows. In Section 2, design of a correlation filter-based optoelectronic tracking system is presented. Computer simulation results are given and discussed in Section 3. Finally, our conclusions are summarized in

2. Design of tracking system

2.1. Problem formulation

One-dimensional notation is used for simplicity. Let us consider a discrete sequence of images where $s_i(x)$ denotes an *i*-th image. Let $t(x;\theta)$ denote a target. θ is a vector of parameters that determine the appearance (distortion) of the target, such as rotation or scaling. Let x_i and θ_i denote the position and the appearance of the target in the *i*-th image, respectively. The problem consists of calculating an estimation $\hat{x}_i$ of the target position in the *i*-th image. No assumption is made on the relation between x_i and x_{i+1}. Location estimation is performed using a bank of composite filters described by transfer functions $\{H_j(\omega), j = 1,2,...N_h\}$ where N_h is the number of filters in the bank. Each filter $H_j(\omega)$ is composed using training images of different views of the target.

2.2. Optimum filter for single-frame detection

In this section we derive an optimum filter for detecting a target in a single frame using a noisy view image. Let $r(x)$ denote a reference image showing one view of the target and $w(x)$ denote the shape of the target in $r(x)$. $w(x)$ takes a value of unity inside the target area and zero otherwise. The suffix is omitted in this section since only one frame is used for the design of optimum filters. So, $s(x)$ is an observed scene represented by a nonoverlapping signal model. The target is opaque and appears over a background that is spatially disjoint. Additive noise is also considered to be present the frame. Filters will be designed from a signal model in which a reference image is corrupted by additive noise. This can model the case when the reference image is captured in controlled environment with low quality equipment and no processing is done on the captured images. The model is formally defined as