Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Robust multi-scale ship tracking via multiple compressed features fusion

Q1 Fei Teng, Qing Liu

Q2 *School of Energy and Power Engineering, Wuhan University of Technology, Wuhan 430070, China*

ARTICLE INFO

Article history:
Received 15 June 2014
Received in revised form
12 December 2014
Accepted 15 December 2014

Keywords:
Compressive sensing theory
Random measurement matrices
Naive Bayes classifier
Coarse-to-fine strategy

ABSTRACT

In this paper, we address the problem of tracking a single ship in inland waterway closed circuit television (CCTV) video sequences given its location in the first frame and no other prior information. First, based on the compressive sensing theory, we employ two kinds of random measurement matrices to extract two complementary good features to track the target ship. Second, in order to track both location and scale, we construct our random measurement matrices according to spatial and temporal structure constraints in consecutive frames, which can be easily obtained and recorded in an offline manner. Having obtained the low-dimensional features in the compressed domain, we further take the different discriminability strengths of the extracted features into account and perform feature evaluations through their cumulative classification performances. A naive Bayes classifier with online update is employed to determine whether the image patch belongs to the foreground or background and a coarse-to-fine strategy is adopted to speed up the time-consuming detection procedure. Finally, both qualitative and quantitative evaluations on numerous challenging CCTV videos demonstrate that the proposed algorithm outperforms Q3 several state-of-the-art methods in terms of accuracy, precision and robustness

1. Introduction

The intelligent video analysis in inland waterway CCTV video sequences has got increasing attention in the past few years [1]. This task usually involves three main steps: detecting the moving ships, tracking the target ship from frame to frame, and analyzing the tracked ship to recognize its behavior [2]. Therefore, ship tracking is a prerequisite in CCTV video surveillance.

In the complex scene of inland waterway, performing long-term robust ship tracking is very challenging due to factors such as pose and illumination change, partial or full occlusion, abrupt scale variation and motion blur, just to name a few. Grabner et al. [3] proposed an online boosting feature extraction method (OAB) for robust tracking. Since the appearance model is constructed and updated with limited (only one positive sample, i.e., the current tracker location) and potentially misaligned samples, OAB is easy to encounter the tracking drift problem. In order to alleviate the tracking drift problem, in [4], a semi-supervised boosting method (SemiB) has been presented in which only the samples in the first frame are labeled and all the other samples are unlabeled. The Multiple Instance Learning (MIL) tracker [5] is also proposed to avoid tracking drift problem. This seminal work uses an MIL based appearance model to represent training data in the form of bags. It rules that a positive bag should contain at least one positive example and examples in negative bag are all negative. The classifier is then trained in an online manner using the bag likelihood function. The reliable experimental results demonstrate that using MIL instead of traditional supervised learning can avoid drift problem and can therefore lead to a more robust tracker with fewer parameter tweaks [6]. However, we observe that online MIL boosting approach is easy to select the less discriminative

---

E-mail addresses: changeerhao_love@126.com (F. Teng), qliu200@163.com (Q. Liu).

http://dx.doi.org/10.1016/j.image.2014.12.006
0923-5965/

features which can confuse the classifier. Moreover, MIL tracker has the potential overfitting problem. In order to solve the potential overfitting problem, Teng et al. [7] proposed Enhanced Multiple Instance Learning (EMIL) method that uses a dynamic function to estimate probability of instance instead of the original logistic function. Kalal et al. [8] proposed a novel Tracking-Learning-Detection (TLD) framework that explicitly decomposes the long-term tracking task into tracking, learning and detection. The tracker recursively follows the object of interest from frame to frame. The detector locates the most likely candidates of image patches. The learning estimates the error of the detector and avoids these errors with online updating P-expert and N-expert. The overall TLD system is robust to non-rigid deformation and scale variation, but is very sensitive to illumination change and cluttered background [9]. Recently, sparse coding [10–13] has been successfully applied in visual tracking with promising results. Zhang et al. [14] proposed the compressive tracking (CT) method. They employed a sparse non-adaptive measurement matrix to extract the low-dimensional features from a multi-scale image feature space with data-independent basis. Then the compressed features are classified via a naive Bayes classifier with online update. This method achieves favorably high efficiency and accuracy in inland waterway ship tracking. However, we observe that CT tracker does not perform well when the ship suffers illumination change and texture variation. Fast compressive tracking (FCT) [15] is an extension to CT that adopts a coarse-to-fine search strategy to reduce the computational complexity in the detection procedure. However, in both CT and FCT, multi-scale tracking for object of interest is not considered. In order to track the scale, Wu et al. [16] develop a MSCT tracker based on normalized rectangle features extracted in the adaptive compressed domain into the bootstrap filter and utilize the target velocity to estimate the current scale. Teng et al. [17] proposed a multi-scale ship tracking method via random projections (MSRP). They develop their ship appearance model based on fern features in the compressed domain in order to track in scenes with light and texture changing. Then, they track the scale by enhancing the tracker with a mechanism of feedback. The extensive qualitative and quantitative evaluations on numerous challenging CCTV videos demonstrate that MSRP can track the scale well, but not stable enough when the ship undergoes extreme illumination change and cluttered background.

Motivated by the above-mentioned discussions, in this paper, we propose an effective and efficient ship tracking algorithm based on hybrid feature extraction and feature selection method. The remainder of this paper is organized as follows: Section 2 describes our proposed method in details. Section 3 discusses the differences with related work. Section 4 conducts numerous experiments on challenging CCTV video sequences to illustrate the effectiveness of the proposed method. Finally, Section 5 presents a brief summary.

2. Proposed algorithm

It is generally well accepted that a typical tracking system should consist of three main components [18]: an appearance model which evaluates the likelihood that the object of interest is at some particular location, a motion model which relates the locations of the object over time and a search strategy for finding the most likely object location in the current frame. We introduce our proposed tracker according to these three components sequentially (Fig. 1).

2.1. Appearance model

In the domain of inland waterway CCTV ship tracking, the target ship may suffer dramatic appearance change caused by pose variation, illumination change, partial or even full occlusion and extreme scale change, which makes the task very challenging. Therefore, an adaptive appearance model is of critical importance for performing long-term robust ship tracking. An appearance model is composed of object representation and statistical model. These two components in our work are described in Sections 2.1.1 and 2.1.2 in details.

2.1.1. Object representation

This subsection aims to deal with the issue of how to use different types of visual features to design a robust ship descriptor. In our work, for each input image patch $\boldsymbol{I} \in \mathbb{R}^{w*h}$, its multi-scale representation is constructed by convolving $\boldsymbol{I}$ with a set of rectangle filters at multiple scales $\{\boldsymbol{H}_{1,1}, \ldots, \boldsymbol{H}_{w,h}\}$ defined as:

$$h_{p,q}(x,y) = \begin{cases} 1, & x_i \le x \le x_i + p,\ y_i \le y \le y_i + q \\ 0, & otherwise \end{cases}. \tag{1}$$

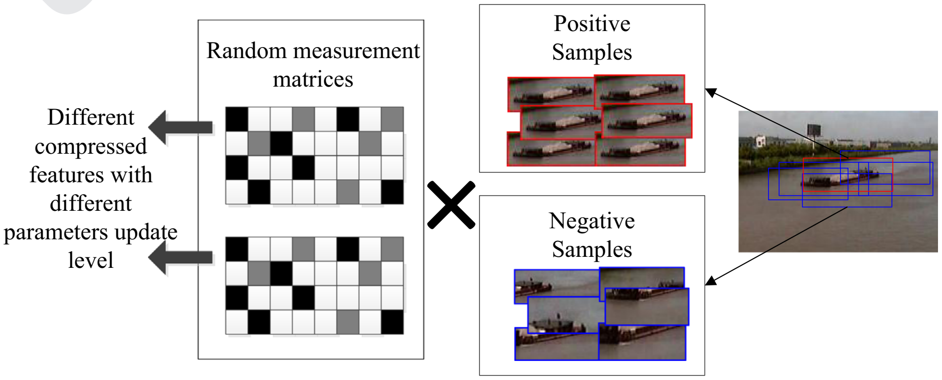

Fig. 1. Feature extraction and parameters update process racker according to these three components sequentially.

Please cite this article as: F. Teng, Q. Liu, Robust multi-scale ship tracking via multiple compressed features fusion, Signal Processing-Image Communication (2014), http://dx.doi.org/10.1016/j.image.2014.12.006