Infrared Physics & Technology 74 (2016) 38–43

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

An object tracking method based on guided filter for night fusion image

Xiaoyan Qian *, Yuedong Wang, Lei Han

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, 210016 Nanjing, China

HIGHLIGHTS

- Automatic guided image filter is applied into fusion image tracking.
- Simple boundary extracted scheme and model updating make the track more accurate and robust.
- This method is validated through subjective and objective metrics.

ARTICLE INFO

Article history:
Received 25 June 2015
Available online 17 November 2015

Keywords:
Target tracking
Guided filter
Color fusion image
Observation model

ABSTRACT

Online object tracking is a challenging problem as it entails learning an effective model to account for appearance change caused by intrinsic and extrinsic factors. In this paper, we propose a novel online object tracking with guided image filter for accurate and robust night fusion image tracking. Firstly, frame difference is applied to produce the coarse target, which helps to generate observation models. Under the restriction of these models and local source image, guided filter generates sufficient and accurate foreground target. Then accurate boundaries of the target can be extracted from detection results. Finally timely updating for observation models help to avoid tracking shift. Both qualitative and quantitative evaluations on challenging image sequences demonstrate that the proposed tracking algorithm performs favorably against several state-of-art methods.

1. Introduction

Object tracking is a fundamental problem with wide application and a rich literature. It is significant for applications such as airport security, construction site safety, patient safety and hospital asset management [1–3]. However, variable lighting, quick moving and random occlusions present difficulties for real-time surveillance, which tend to cause erroneous object detection and trajectories. Wang et al. proposed that combining the advantages of different visual sensors can improve the robustness of object detection and tracking and this is a promising field [4]. Image fusion aims to integrate multiple images derived from different sensors into a composite image that is more suitable for the purposes of human visual perception or computer-processing tasks [5].

Recently, more and more researchers are interested in using information from different sensors to improve the tracking performance [1,5–10]. They developed pixel-, feature- and decision-level fusion techniques for tracking. Pixel-level method works in the spatial domain or in the transform domain. Image fusion at pixel level amounts to integration of low-level information and the fused image can maintain the rich color of visual image and meanwhile pop out the target in the infrared image [1]. Some researchers improved the tracking performance using the appearance features from the fused results [6]. Qian et al. [5] proposed an object observation model based on 2D color histogram and gave the local matting tracking scheme for the night fused images. The feature-level algorithms typically calculate features from each image and fuse their properties to realize object tracking [4,7]. Zhao et al. [7] proposed an object tracking method based on infrared and visible dual-channel video. It extracted the Hue, Saturation and value color features in visible image and used mean shift to estimate the object location in the visible image. The contour feature in infrared image realized accurate tracking [8,9]. Liu et al. fused tracking in color and infrared images using joint sparse representation [9]. In this paper, a similarity induced by joint sparse representation is designed to construct the likelihood function of particle filter tracker so that the color visual spectrum and thermal spectrum images can be fuse for object tracking. Decision-level methods get the tracking results separately in each source image and then use the outputs of initial object tracking as inputs to fusion algorithm to produce the object state. For example, Conaire et al. [10] proposed a framework that can efficiently combine features for robust tracking based on fusing the outputs of multiple spatiogram trackers. The framework allows the features to be split

* Corresponding author.

http://dx.doi.org/10.1016/j.infrared.2015.11.005
1350-4495/

arbitrarily between the trackers, as well as providing the flexibility to add, remove or dynamically weight features.

In this paper, we focus on object tracking for pixel-level fused images. A tracking method typically consists of three components: an observation model (e.g. contours [11] and histograms of oriented gradients [12,13]), which evaluates the likelihood of an observed image patch belonging to the object class; a dynamic model which aims to describe the states of an object over time (e.g., Kalman filter [14] and particle filter [15,16]); and a search strategy for finding the likely states in the current frame (e.g. sliding window [17] and mean shift [18]). In this paper, we propose a robust algorithm using a generative appearance model that considers the effects of occlusion to alleviate tracking drift.

In order to develop effective observation models for object tracking, early works tended to construct the model by describing the target itself [19], while recently the adoption of context information has become very popular [20]. Although these methods more or less suffer from inaccuracy in the estimation of foreground and background which will cause tracking shift, they motivate us to design an accurate observation model which can capture the information from the target and its neighboring context. To focus on the alleviation of the aftereffects caused by foreground/background labeling errors, an accurate boundary of the target could mostly reduce such errors. There are several tracking methods that try to obtain boundary of the foreground [5,21,22]. Tracking using active contours is one way to extract and track the object boundaries [23]. However, active contour tracking heavily relies on curve matching and is not designed for complex shape deformation; therefore, it cannot handle large deformation. Image segmentation is another direct and popular solution to separate the foreground from the background [24]. In [21], a shape-based method was proposed to match the contour of a prior shape to the current image, but contour matching alone may not achieve good results since the useful information within the target boundary is not taken into account. In [5], a matting tracking method was proposed. Compared with image segmentation, matting can exploit the linear compositing equations in the alpha channel instead of directly handling the complexity in color image. Therefore, it may achieve better foreground/background separation performance. In addition, the adaptive appearance model automatically generates scribbles in each frame, which makes the performance of foreground/background separation only rely on the scribble. Their model adaptation largely excludes the ambiguity of foreground and background, thus these methods significantly alleviate the drift problem in tracking.

Benefiting from the matting tracking, we propose a robust generative tracking algorithm based on guided image filter in this paper. Different from matting tracking, the detection of foreground target can be produced by filtering a raw appearance scribble under the guidance of the source image. There are three advantages for guided image filter to be applied into tracking. One is that the time complexity is independent of the window radius, which allows us to select random kernel size; the second is that it avoids solving the matting Laplacian matrix, which will bring faster tracking. The last one is that this filter can well maintain the guidance's edge information, which allows us to get real contour. So object scaling and rotation can be handled by obtaining the accurate and robust boundary. In addition, our discriminative model adaptation largely excludes the ambiguity of foreground and background, thus significantly alleviating the drift problem in tracking. The adaptive appearance model can handle partial occlusion and other challenging factors. Experiments and evaluations on fusion image sequences bear out that the proposed algorithm is efficient and effective for robust object tracking. Fig. 1 gives the flow chart of our algorithm.

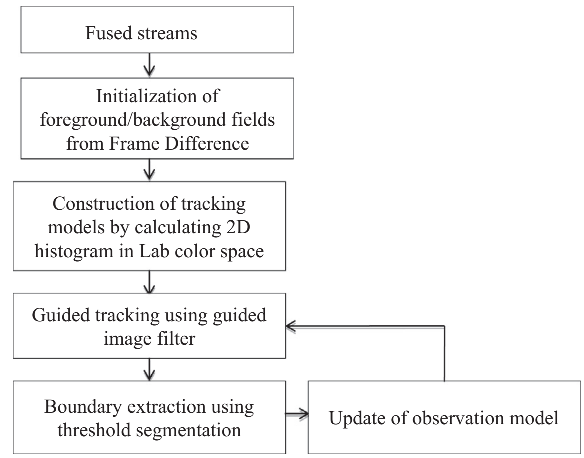

Fig. 1. The flow chart of our algorithm.

2. Guided image filter

Guided image filter is an edge-preserving smoothing filter. It avoids the gradient reversal artifacts that may appear in detail enhancement. The key assumption of the guided filter is a local linear model between the guidance I and the filter output O:

$$O_i = a_k I_i + b_k, \ \forall i \in w_k \tag{1}$$

where (a_k, b_k) are some linear coefficients assumed to be constant in a window w_k centered at the pixel k. This local linear model ensure that O has an edge only if I has an edge, because $\nabla O = a \nabla I$. By minimizing the difference between O and the filter input P, the linear coefficients are determined by:

$$\begin{aligned} a_k &= \frac{\frac{1}{|w|}\sum_{i \in w_k} I_i P_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon} \\ b_k &= \bar{p}_k - a_k \mu_k \end{aligned} \tag{2}$$

Here, μ_k and σ_k^2 are the mean and variance of I in w_k, $|w|$ is the number of pixels in w_k, and $\bar{p}_k = \frac{1}{|w|}\sum_{i \in w_k} P_i$ is the mean of P in w_k.

The relationship among I, P and O can be rewritten in the form of image filter like the following:

$$O_i = \sum_j W_{ij}(I) P_j \tag{3}$$

Here, $W_{ij}(I)$ is the kernel weight which can be explicitly expressed by:

$$W_{ij}(I) = \frac{1}{|w|^2} \sum_{k:(i,j) \in w_k} \left(1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon}\right) \tag{4}$$

Compared with the closed-form solution to matting [22], we find that the elements of the matting Laplacian matrix L can be directly given by the guided filter kernel weight:

$$L_{ij} = |w|(\delta_{ij} - W_{ij}) \tag{5}$$

where δ_{ij} is the Kronecker delta. So if there is a reasonably good guess of the matte, we can run guided filtering process to produce a fine alpha matte just like Laplacian matting. By simple boundary extraction, the opacity map can give accurate and whole detection of tracking object.

Just like matting tracking, we also can fit guided filter naturally into the tracking framework by filling the gaps including constructing the appearance model, giving the search strategy, generating matte automatically from the model and performing model updating.