



Robust infrared target tracking using discriminative and generative approaches



C.S. Asha*, A.V. Narasimhadhan

National Institute of Technology Karnataka, Surathkal, Mangalore 575025, India

HIGHLIGHTS

- This paper presents a gradient and channel coded features under kernelized correlation filters.
- A novel approach to combine complementary techniques for drift free tracking.
- An efficient scale estimation of target is proposed.

ARTICLE INFO

Article history:

Received 21 January 2017

Revised 1 May 2017

Accepted 31 May 2017

Keywords:

Correlation filter
Channel coded feature maps
Multi frame template
Adaboost classifier
LK scale estimation

ABSTRACT

The process of designing an efficient tracker for thermal infrared imagery is one of the most challenging tasks in computer vision. Although a lot of advancement has been achieved in RGB videos over the decades, textureless and colorless properties of objects in thermal imagery pose hard constraints in the design of an efficient tracker. Tracking of an object using a single feature or a technique often fails to achieve greater accuracy. Here, we propose an effective method to track an object in infrared imagery based on a combination of discriminative and generative approaches. The discriminative technique makes use of two complementary methods such as kernelized correlation filter with spatial feature and AdaBoost classifier with pixel intensity features to operate in parallel. After obtaining optimized locations through discriminative approaches, the generative technique is applied to determine the best target location using a linear search method. Unlike the baseline algorithms, the proposed method estimates the scale of the target by Lucas-Kanade homography estimation. To evaluate the proposed method, extensive experiments are conducted on 17 challenging infrared image sequences obtained from LTIR dataset and a significant improvement of mean distance precision and mean overlap precision is accomplished as compared with the existing trackers. Further, a quantitative and qualitative assessment of the proposed approach with the state-of-the-art trackers is illustrated to clearly demonstrate an overall increase in performance.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The human eye is sensitive to visible light that comprises a small portion of the electromagnetic spectrum. The object reflects, absorbs or transmits energy depending upon its properties. The human eye absorbs the light reflected from the object allowing the object to be viewed as it is. Visual cameras require an illumination source to capture images through reflected light. On the other hand, thermal infrared (TIR) cameras do not rely on a lighting source as they acquire the image based on the temperature that is radiated from the object. Infrared technology was initially employed in military applications and has subsequently migrated

to various scientific and medical areas. The night vision camera, a relevant coinage of infrared technology is extensively utilized in intricate mechanisms such as building inspection, gas detection, complex industrial procedures, medical, veterinary, agricultural, fire detection, surveillance, aerospace, target acquisition and tracking of humans, vehicles and animals [1]. Thus, the power of thermal cameras to capture images under all climate conditions, create a larger impact on its applications.

Infrared (IR) is an invisible electromagnetic spectrum with longer wavelengths than the visible spectrum. Based on wavelength, the infrared band is subdivided into different ranges as shown in Fig. 1. The infrared range includes near infrared (0.7–1 μm wavelength), short wave infrared (1–3 μm wavelength), mid wave infrared (3–5 μm) and longwave infrared (8–14 μm). Unlike visual or near infrared cameras that capture radiation

* Corresponding author.

E-mail address: asha.cs@rediffmail.com (C.S. Asha).

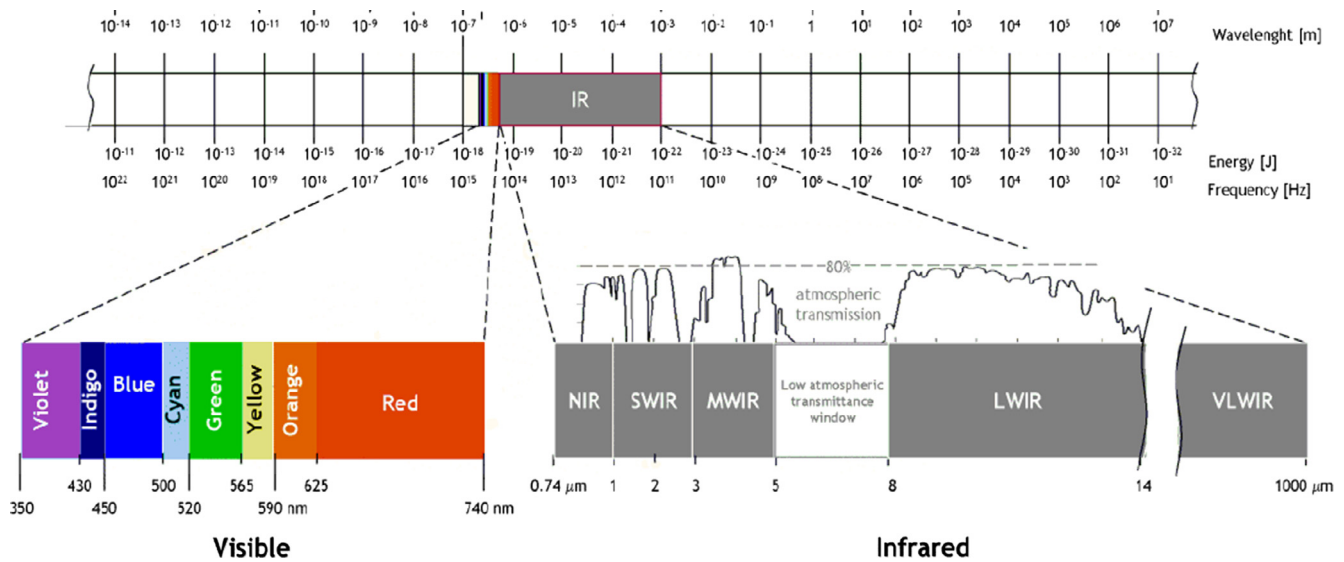


Fig. 1. infrared and visible spectrum.¹

reflected by objects, long wave or mid wave infrared cameras capture radiation emitted by objects.

The objective of infrared tracking is to find the target trajectory over subsequent frames beginning from the initial ground truth location in the first frame. Compared to RGB recordings, infrared tracking is additionally challenging due to the following reasons [2]:

- Thermal images are noisy by nature (i.e. low signal to noise ratio (SNR) characteristics), they are of poor quality and low resolution, containing large numbers of dead pixels.
- Thermal infrared images appear as gray-scale images with missing texture and color highlights.
- The intensity of objects in thermal infrared images varies based on temperature as opposed to illumination.
- Handling occlusion and re-identification of objects in thermal infrared imagery is as difficult as separating two objects of comparable shape and color. For instance, individuals strolling in a group, animals of the same sort, and fowls in a rush share similar shape and intensity levels in thermal imagery.
- For more sophisticated backgrounds, it is difficult to detect and track the object because the target of interest may blend with the background, or it may change in size, shape or intensity.

In summary, It is difficult to determine the unique properties of the target in thermal imagery.

Tracking algorithms mainly differ in terms of usage of object features, motion model, memory and object representation. The algorithms that have been proposed for RGB videos can be described as generative and discriminative approaches. The generative approaches build an appearance model to describe the object and search locally for the most similar region in every frame. Such techniques typically employ template matching using normalized cross correlation (NCC) [3], histogram matching by maximizing similarity function [4], subspace techniques such as principal component analysis to represent object appearances [5], spatial histogram features to minimize l_1 distance between object model and target in Distribution Field Tracker (DFT) [6], channel coded features to model the object and minimizes l_1 distance between object model and target in Enhanced Distribution Field Tracker (EDFT) [7]. However, due to the local search method employed in generative techniques, these methods normally fail when different

objects with similar properties are encountered, which is a common problem in infrared images.

The discriminative approaches are based on target by detection, where tracking is treated as a binary classification task. The objective of classifier-based approaches is to determine the decision boundary in order to distinguish the object from the background without building a complex model of the object. Numerous detection and classification techniques have been produced over time, using appropriate features and suitable classifiers. For efficient target tracking, Grabner et al. employed AdaBoost classifier using haar, oriented histogram and local binary pattern features of the object [8]. In Fast Compressive Tracking (FCT) [9], haar features of the object and background are exploited to train the Naive Bayes classifier. However, the local search method used in discriminative techniques result in drifting of the target in the case of similar objects. Among the several discriminative types of trackers, correlation filters make predominant contributions in tracking, due to high frame rate and efficiency. Minimum Output Sum of Squared Error (MOSSE) [10] filter achieved improved tracking performance by employing gray scale feature in the correlation filter. As an enhancement, Kernelized Correlation Filters (KCF) have been proposed by exploiting circulant data structure for training and testing using histogram of gradient (HOG) channel features [11]. Under linear correlation filter framework, Danelljan et al. used HOG features with accurate scale estimation in DSST [12]. In distractor aware tracking (DAT) [13], the color statistic model is utilized and distracting regions are identified in advance to avoid drifting.

A limited number of methods have been proposed for infrared tracking, that include pedestrians or small dim objects. By utilizing intensity and edge features under particle filter, pedestrian tracking has been proposed [14]. The Gaussian Mixture Model (GMM) has been employed to separate the foreground from the background [15]. In addition, the shape-based features are trained with a support vector machine (SVM) classifier to detect pedestrians. Post detection, intensity combined with edge cues under particle filter are used to track pedestrians in infrared videos. The tracker presented in [16], exploits motion prediction techniques to detect any false alarm and activates template matching for object recovery. A combined technique based on curve matching and Kalman filter has been proposed in [17] to predict target position in infra-

¹ Source: www.mivim.gel.ulaval.ca.

Download English Version:

<https://daneshyari.com/en/article/5488593>

Download Persian Version:

<https://daneshyari.com/article/5488593>

[Daneshyari.com](https://daneshyari.com)