Object Tracking in Surveillance System using Gaussian-Sum Filter And ACF Detection

Christanto Febrian Wibisono
School of Computing
Telkom University
Bandung, Indonesia
febrianwibisono@student.telkomuniversity.ac.id

Hilal Huda
School of Computing
CAATIS, Tel-U
Bandung, Indonesia
hilalh@teluapp.org

Muhamad Irsan
School of Computing
Telkom University
Bandung, Indonesia
irsanfaiz@telkomuniversity.ac.id

Aji Gautama Putrada
School of Computing
Telkom University Jakarta
Jakarta, Indonesia
ajigps@telkomuniversity.ac.id

Syifak Binti Izhar Hisham
Science Computer
Universiti Malaysia Pahang Al-Sultan Abdullah
Pekan, Malaysia
syifakizhar@umpsa.edu.my

***Abstract*— This paper presents a study on object tracking in surveillance systems using Gaussian-Sum Filter and Aggregate Channel Features (ACF) detection to address the challenges of accurately tracking multiple objects in dynamic environments. Object tracking is crucial in computer vision, with applications from surveillance and security to autonomous navigation and robotics. This study employs the Gaussian-Sum Filter, a proven Bayesian filtering algorithm known for its predominance in non-linear scenarios, which keeps object tracking more consistent over time. However, since the ACF detection method can detect objects over multiple frames with higher accuracy than our initial detections, we combine it with initial ones. Performance testing is conducted across four datasets, using key metrics such as precision, Multiple Object Tracking Precision (MOTP), and Multiple Object Tracking Accuracy (MOTA) to evaluate effectiveness. The results show that while Gaussian-Sum Filter combined with ACF detection achieves different precision with specific datasets (7%-98%) and MOTP rates (10%-73%), challenges arise in maintaining uninterrupted tracking accuracy, as evidenced by very low MOTA (-6%-10%) and a significant rate of false negatives, especially in complex scenarios with occlusions. These findings suggest that although Gaussian-Sum Filter and ACF detection are effective for initial detection and data handling, enhancements or hybrid methods may be required for applications demanding high accuracy in continuous multi-object tracking.**

Keywords— Object Tracking, Multi-Object Tracking (MOT), Gaussian-Sum Filter (GSF), Aggregate Channel Features (ACF), Computer Vision

I. Introduction

A. Background

Object tracking is an important task in computer vision that focuses on detecting and following objects in a series of images. Object tracking applicable in areas such as traffic monitoring, video surveillance, and digital city infrastructure [1]. Object tracking is usually divided into two main types based on the task. Single Object Tracking (SOT) is used to follow one specific target in a video, while Multiple Object Tracking (MOT) or Multiple Target Tracking (MTT) is used to track several objects at the same time [2].

Several studies have revealed that object tracking is essential in computer vision. One study showed that tracking moving objects in video image sequences is important in computer vision. This study also focused on object tracking, which has already been used in many areas of computer vision, including video surveillance, artificial intelligence, military guidance, safety detection, robot navigation, and medical and biological applications. [3]. Han *et al.* [4] also explain that Object tracking is the process of detecting and following objects as they move through video frames which involve identifying objects predicting their movement and continuously updating their position over time while addressing challenges like occlusion changes in appearance and complex movements to ensure accurate and continuous tracking of objects in dynamic environments such as surveillance or sports. Furthermore, other studies have investigated among the various nonlinear estimators, the Gaussian Sum Filter (GSF) stands out as a flexible and effective method for addressing nonlinear estimation problems, it provides a robust framework that can adapt to different types of estimation tasks, making it a valuable tool in many fields, over time, the GSF has been extensively studied and discussed in numerous research papers, highlighting its importance and versatility in solving complex estimation challenges [5]. In the context of Multiple Object Tracking Luo *et al.* [6] explain Multiple Object Tracking (MOT) or something that is more frequently referred to as Multiple Target Tracking (MTT), enables computer vision systems to detect the location of multiple objects in a scene while also re-identifying each object across frames to produce their trajectories over the sequence, this process plays a critical component in any kind of surveillance and tracking tasks. In data association-based MOT, the tracking performance is heavily affected by the detection results [7].

Gaussian filter is a Bayesian technique for state estimation in systems characterized by Gaussian-distributed states and noise. By updating the state estimate using Bayesian principles and new observations, it provides an efficient approach for systems with linear dynamics, as seen in the Kalman Filter and its variants. Dore *et al.* [8] Bayesian approach in video tracking offers a robust framework for managing uncertainties and improving object tracking accuracy by integrating motion dynamics and noisy observations, with recursive updates that enable adaptation to change in appearance, occlusions, and other common video analysis challenges.

Gaussian-Sum Filter is a type of Bayesian filter that estimates the state distribution using a weighted combination of Gaussian probability density functions. This allows for more accurate approximations in nonlinear systems. [9]. Terejanu *et al.* [10] said that The Gaussian Sum Filter (GSF) has seen various improvements, such as advanced measurement updates, mixtures of Kalman filters, and Gaussian sum particle filtering, with applications in fields like target tracking, computer vision, and geoscience. However, in all these methods, the weights of the Gaussian components remain fixed between measurements, limiting the accuracy of the GSF when measurement data is sparse.

This paper proposes the application of a Gaussian-Sum Filter for Multi-Object Tracking (MOT) in an online, real-time tracking system. The methodology begins with loading a pedestrian video dataset from the MATLAB website into a program that utilizes the Gaussian-Sum Filter. The program generates two types of detections ACF, followed by a performance analysis of the Gaussian-Sum Filter. The analysis evaluates key metrics, including MOTA, MOTP, precision, and false negative.

In this section, we outline the structure of the rest of the paper. Section II details the methodology used in this study. Section III presents the performance results of our approach, and Section IV provides the conclusions drawn from the research.

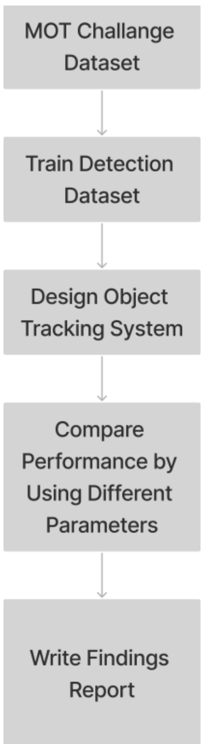

Fig. 1. Research Workflow for Object Tracking Using Gaussian-Sum Filter

II. METHODOLOGY

A. Multi-Object Tracking

Multi-Object Tracking (MOT) is one of the most significant areas in intelligent video surveillance and autonomous vehicles. It allows for the tracking of multi-targets and plays a vital role in many applications, such as traffic management, security monitoring, and self-driving systems, by accurately tracking objects. [11]. Moraffah *et al.* [12] also explain that In the past decade, multiple object tracking (MOT) has been a tough and time-consuming problem, used in many areas like computer vision, driver assistance systems, surveillance, and radar tracking, where it is important to track and follow several objects accurately at the same time.

According to Bernardin *et al.*[13] Multi-Object Tracking (MOT) is a technique used to monitor multiple objects simultaneously within a sequence of frames, often applied in video surveillance and motion analysis, two key metrics are used: MOTA (Multiple Object Tracking Accuracy) and MOTP (Multiple Object Tracking Precision). MOTA measures the system's accuracy in maintaining object identities, considering errors like false positives, missed detections, and ID switches. Two very intuitive metrics can be defined:

$$MOTP = \frac{\sum_{i,t} d_{i,t}}{\sum_t c_t} \tag{1}$$

Where $d_{i,t}$ represents the distance between the estimated position and the ground truth position for a tracked object i at time t, and c_t is the total number of matches (or correctly detected objects) in each frame t.

Multiple Object Tracking Precision measures the spatial accuracy of tracking by calculating the average distance between the predicted and actual positions of tracked objects. It indicates how precisely the tracker estimates each object's location within the frame. The formula for MOTP is:

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \tag{2}$$

where m_t, fp_t and mme_t are the number of misses, of false positives and of mismatches respectively for time t. The MOTA can be seen as composed of 3 error ratios:

$$\overline{m} = \frac{\sum_t m_t}{\sum_t g_t}, \tag{3}$$

the ratio of misses in the sequence, computed over the total number of objects present in all frames,

$$\overline{fp} = \frac{\sum_t fp_t}{\sum_t g_t}, \tag{4}$$

the ratio of false positives, and

$$\overline{mme} = \frac{\sum_t mme_t}{\sum_t g_t}, \tag{5}$$

the ratio of mismatches.

MOTA measures all the errors in object configurations made by the tracker, including false positives, missed detections, and mismatches across all frames, offering a clear assessment of tracking accuracy.

B. Gaussian-Sum Filter

The Gaussian-sum approximation has been introduced and proposed as an effective method to achieve practical implementation of nonlinear Bayesian filtering, providing a way to handle the complexities involved in nonlinear systems