

**Complex Systems**

**(6CS014)**

**Title: Object Recognition and Tracking**

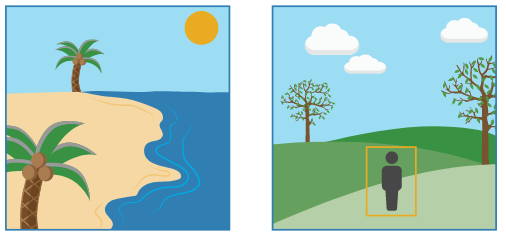
|  |  |
| --- | --- |
| Student Id | : 2049822 |
| Student Name | : Sashank Dulal |
| Supervisor | : Mr. Rupak Koirala |
|  |  |
|  |  |
| Submitted on | : Wednesday, September 16,2021 |

# Acknowledgements

# Table of Contents

# Introduction

Object detection is a technique used in computer vision to identify the objects within an image or video. The main objective of this method is to be able to replicate the recognition abilities of humans with the use of various processing techniques and methods (mathworks, Inc, 2021). The technologies “object detection” and “object recognition” are similar in terms of their objectives but the major difference is that the object is identified as well as located in the image using the object detection technique (mathworks, Inc, 2021).



Similarly, object tracking is another image processing technique that involves detection of object, creation of unique identifier for each object in the video and tracking of the object as they are moving around different parts of a frame in a video (viso.ai, 2021).

# Aims and Objectives

## Aims

* Gain an in-depth understanding of object detection and tracking.
* Research on various techniques available for detection and tracking of objects.

## Objectives

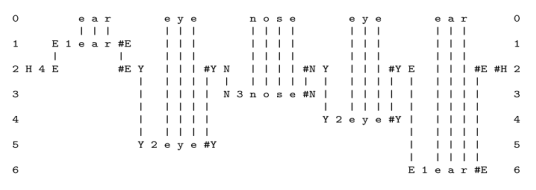
* Research on similar systems and published literatures.
* Analyze their qualities and drawbacks.
* Research on the complexity of the techniques available.

# Literature Review

## Detection, Recognition and Tracking of Moving Objects from Real-time Video via SP Theory of Intelligence and Species Inspired PSO

This paper mainly focuses on analyzing the SP Theory of Intelligence and address the challenges in tracking moving objects. In this paper, the authors have extracted multiple alignments, parsed raw data to gather noise-free alignment and reach best-possible solution with the user of family resemblance concept in order to efficiently recognize the object. Then the recognized object is tracked using Particle Swarm Optimization (PSO) inspired by species.

In this paper, the authors have proposed using SP theory of intelligence to detect objects which is mainly efficient in NLP due to its ability to process noisy and erroneous data and after the detection of objects have been done, species-inspired Particle Swarm Optimization (PSO) was to be used to track the objects. The use of SP system to detect the object can be demonstrated by the image below in which depicts how a person’s face with all the features are parsed.



Then the knowledge of the structure of features (Nose, Ears etc) are aligned with every feature of the object of interest. Then the alignment is checked with the pattern in row 2 (i.e. comparing for head). Because of this process, much of the errors due to noisy data is eliminated.

Now for the tracking of objects, a PSO framework is used where

(Ray, Dutta, & Chakraborty, 2017)

## Object Recognition and Tracking for Remote Video Surveillance

This paper proposes a real-time object recognition and tracking system for surveillance. In order to meet the real-time requirement, the authors have come up with the idea of using statistical morphological skeleton to achieve accuracy in localization, low complexity in terms of computation and more strength against noise. Recognition of object is achieved by the comparison of the approximation of the skeleton function obtained from the analysis of image with the one obtained from the model store in the database and the tracking task is performed by applying “Kalman filter” to the set of quantities observable gathered from the skeleton and with the help of other characteristics of the moving object.

Proposed System

In the proposed system, a camera is used to gather video of the surveillance scene. Then a change detection algorithm is applied to every frame of the video which gives out a binary image where each possible object is represented with a blob moving inside the scene. The absolute difference between the current image and a background image is computed at a fixed time interval. The formula to calculate the difference is given below:

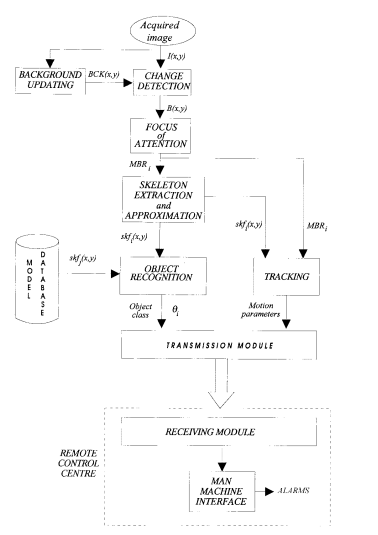


Where D(x,y) is the difference, I(x,y) is the current image and BCK(x,y) is the background image. The resulting image is N \* M in size. Then a hysteresis function having two thresholds, THRin and THRout and two states (i.e. background and object) is applied to the difference image to decide whether a point is either background or a moving object.

With this method, with every calculation, a binary image with changing pixels denoted as 1 and background pixels as 0 is generated. Then the Kalman filter is applied to each pixel of the image to pet the predicted background estimate. The Kalman filter is represented by the following equations:



Where S(i,p) is the level of grayness of the background, S(i+1,p) is the same quantity of grayness in the next frame and µ(i) is an estimated system error. Similarly, l(i,p) is the level of gray in current image point p and η(i,p) represents the estimate of noise in the input image.

Figure 1: Architecture of the proposed system

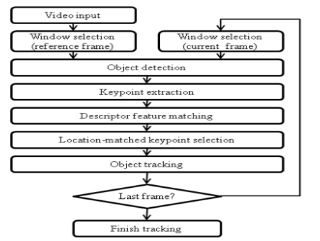
To be continued

(Foresti, 1999)

## Multiple Object Tracking Using SIFT Features and Location Matching

This paper presents a system that is able to rack objects with stable results by removing location-mismatched points. It uses Shift Invariant Transform (SIFT) algorithm which generates local features that are immune to changes in the scale of image, noise, lightings and distortion. In D.G. Lowe’s (who is the author of the SIFT algorithm) proposed paper, the matches having distance ratio greater than 0.8 were rejected so the authors propose a method to increase the efficiency of the algorithm by recognizing location-mismatched points and restraining their participation in the matching of keypoints. Then tracking is applied to the video which determines the size and position of the window.

The proposed algorithm in this paper uses location-matched keypoints from the keypoints generated from SIFT algorithm, the tracking is achieved by rectangle window surrounding the object at the reference frame and the current frame and location-mismatched keypoints are avoided. Firstly, rectangle window surrounds object in reference frame, keypoints are generated through SIFT algorithm and their features are stored. Then, keypoints are genereated in the next frame using same methods, matching keypoints are selected based on the distance ratio and the matched keypoints are selected as the candidate keypoints for tracking. Then the location-matching keypoints are determined which is used for stable tracking of the objects. The overall flow of the algorithm is shown in the image below.



The reference frame and current frame is used to generate keypoints through SIFT algorithm which is as follows:



With this technique, the resulting data are “im” which includes pixel values of the test image, “des” is a matrix of descriptor vectors, and “loc” consists of location, scale and orientation values for every keypoints. The matching of distance in SIFT is mostly done with the help of “des” data. Since all of the keypoints are classified into two parts, the object and the background and the background could change while the object is moving, the keypoint extraction must consider the fact. The decision is taken with the help of formula below:



Remaining

(Ha & Moon, 2011)

## Multi-Object Recognition and Tracking with Automated Image Annotation for Big Data Based Video Surveillance

In this paper, an improved region based scalable convolution neural network (IRS-CNN) based multiple object tracking model has been discussed. It uses Automatic Image Annotation (AIA) in order to improve the detection capacity and reduce computation time.

(Vijiyakumar, Govindasamy, & Akila, 2021)