Tracking Learning and Detection of Multiple Objects using Static Camera

Abhishek Khanolkar, Dhaval Shah, Ayushi Desai

Department of Information Technology
Thakur College of Engineering and Technology
Kandivali East, Mumbai – 400101.

Abstract- Proposed system performs tracking, learning and detection of multiple objects in the video stream using a static camera. Proposed system does the tracking of multiple objects in presence of occlusion, clutter, scaling and rotation. There were many tracking approaches have been proposed in the past decades such as histogram based mean shift methods, contour based tracking methods but they were intensive while achieving better tracking accuracy. Proposed System presents robust tracking framework that can be applied to real time applications. Propose system is expected to track and detect the real time object which is under constant rotation upto 180 degree.

General Terms-Extraction, matching template, tracking, detection.

Keywords motion detection, improved tracking camshift method, real time object learning.

INTRODUCTION

To track moving objects is a key ability for different agents to fulfill many different tasks like surveillance, guiding or following as well as interacting with and learning from humans. Many successful and accurate object tracking approaches have been proposed in recent years. However, many of them are not applicable for the tasks of moving objects, because the domain violates several of the underlying assumptions. There is no static background and no fixed target appearance and the image quality can be bad due to insufficient illumination or glare. In some applications one cannot build a complex target model offline, because the kind of object to track is not known in advance. The purpose of our project Understands objects in video data due to its enhanced automation in public security surveillance as well as in traffic control and pedestrian flow analysis. Here, a system is presented ,This is able to detect and classify people and vehicles outdoors in different weather conditions using a static camera.

PROBLEM DEFINITION

The main focus of our application is surveillance.A recent survey done by the traffic as well as security agency assert that crimes and chaos take place where crowd are there in huge volumes ,thus to keep a track of all the moving as well detect and identify a security system that includes learning detection and tracking of object is developed. Existing system has the problems when object changes its appearances or object is moving out of camera view and again coming in front of camera. Tracking fails due to scaling, rotation, illumination changes and does not perform in case of full out of plane rotation. systems that were in use did not have full functionality regarding identification and tracking of humans as well as objects.These system also lacked tracking of moving objects which resulted in escape of criminals or culprits. Here it is recommended by professionals in the IT world to develop a system inclusive of all learning tracking and detection of objects using a static camera.

DEVELOPMENT IDEA

Camshift approach:
The Camshift algorithm uses histogram technique and iterative mean shift procedure on the back projected generated from new images to locate the image objects location, Camshift automatically adjust the object scale during tracking .A Calculation of the Back-projection Image and locating its center When tracking objects, CamShift uses HSV colour to present the colour distribution, the back projection image is generated from every new frame for processing. In order to reduce the computational consumption during tracking hue channel can be used in future.

Formula
\[ \sum_{x=1}^{n} \delta(c(x)_x-u), (1) \]
\[ Qn=\text{summation} \ d(c(x)-u) \ (2) \]
\[ I=1 \ A \text{ normalized method can be calculated as follow} \]
\[ M00=\sum_{x} \sum_{y} \text{lbp}(x,y) \ (3) \]
\[ \text{The first moment for x and y} : \]
\[ M10=\sum_{x} \sum_{y} x\text{lbp}(x,y) \quad M01=\sum_{x} \sum_{y} y\text{lbp}(x,y) \ (4) \]

The centroid of search window:
\[ Xc= \frac{M10}{M00}, \quad Yc= \frac{M01}{M00} \ (5) \]
2) Analysis of Cam shift algorithm :
Based on the image projection lbp, the process can be described as follows:
An initial local window w is used to find the new center location and set the search window to the 0th moment. Repeat this procedure until an elliptical image can depict the object.
Although the cam shift method is very fast for processing images but it initiates execution from the initial position and converges to local maximum

A Iterative search
We first introduce a constraint that splits search method used by Adaptive local based search.

Algorithm : CamShift-Based Object Tracking Framework with ALS and Kalman Filter
Input: current frame $l_t$, the object window from last frame $W_{t-1}$, similarity threshold T
Output: the object window in the current frame $W_t$
1: calculate the back projection image.
2: Use the kalman filter to predict new location and $l_{ext}$
3: $W_{temp}=A_{cs}(W_{t-1})$
4: calculate similarity $P=P_{cs}(W_{t-1})$
5: if $P>T$ then
6: $W_t=W_{temp}$
7: else
8: $W_t=ALS (W_{t-1}, l_{ext})$
9: end if
10: update the kalman filter
11: return $W_t$

Concept model

Scope of Project
A wide range of possible applications for TLD technology has been identified. Potentially fruitful include:
- Crime prevention: Law enforcement agencies typically maintain large archives of visual evidence, including past suspects, facial photographs and prints. With this data integrated with our software, any recognized person from the database is identified and detected. Surveillance cameras will detect and track him down. Hiding criminals from the system would be impossible.
- Surveillance at schools, malls, and crowded areas: These places where safety is the major concern. The software can be integrated with the camera at these places which can keep a track of the situation and alert surveillance teams during any mishap.
- Traffic signal: Traffic signals are established to maintain law and order, but many people break these laws by abolishing the traffic signal rules. The software can be integrated with a camera and can be established with the signals. The software will identify and track all the people breaking the signal.

Performance Parameters
Based on the matching strategy described above, two very intuitive metrics can be defined.
1) The multiple object tracking precision (MOPT):
$$\text{MOTP} = \frac{\sum_{i=1}^{n} t_i d_{i}^2}{\sum_{i=1}^{n} t_i}$$

2) The multiple object tracking accuracy (MOTA):
$$\text{MOTA} = \frac{1}{\sum_{i=1}^{n} g_i} \sum_{i=1}^{n} (m_i + fpt_i + mm_i)$$

Where $m_i$, $fpt_i$, and $mm_i$ are the number of misses, of false positives, and of mismatches, respectively, for time $t$. The MOTA can be seen as derived from 3 error ratios:
$$\overline{m} = \frac{\sum_{i=1}^{n} m_i}{\sum_{i=1}^{n} t_i}$$
$$\overline{fpt} = \frac{\sum_{i=1}^{n} fpt_i}{\sum_{i=1}^{n} t_i}$$

Performance table

<table>
<thead>
<tr>
<th>FPS</th>
<th>$\alpha$</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>25</td>
<td>65%</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>75%</td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>45%</td>
</tr>
<tr>
<td>27</td>
<td>30</td>
<td>60%</td>
</tr>
</tbody>
</table>

Table based on proposed system
MOT P = multiple object tracking precision
MOT A = multiple object tracking accuracy
FPS = frames per sec
$\alpha$ = (threshold) tracking accuracy with respect to distance
EXPERIMENTAL RESULTS
We implement the motion detector in .net framework using c#. It detects the motion of the objects when the object is moving. We implement 4 motion detectors that detects different background, foreground objects. It has 1 optimized detector for achieving the accuracy in detection.

DISCUSSION
In this project, we studied the problem of tracking of an unknown object in a video stream, where the object changes appearance frequently moves in and out of the camera view. We designed a new framework that decomposes the tasks into three components: tracking, learning, and detection. The learning component was analyzed in detail. We have demonstrated that an object detector can be trained from a single example and an unlabeled video stream using the following strategy: 1) evaluate the detector, 2) estimate its errors by a pair of experts, and 3) update the classifier. Each expert is focused on identification of particular type of the classifier error and is allowed to make errors itself.

CONCLUSION
The system which we are going to implement is based on tracking and identification of multiple real objects. improves the detection and classification of objects. The system can handle occlusions and has demonstrated good results over multiple objects in varying conditions. In each and every case, the system accurately labeled the dynamic objects tracked correct.

ACKNOWLEDGMENTS
I would like to express my deep gratitude to DR. B.K Mishra, Principal, Thakur College of Engineering, and Technology, Mumbai for extending me the opportunity for major project and providing all the necessary, resources for this purpose. I express heartfelt thanks to Mr. Bhushan Nemade for his wonderful support for preparing the project and for giving me an opportunity to do my project on “Tracking Learning and Detection of multiple objects using Static Camera” Relationship Management System”. I am grateful to Mr. Vinayak Bharadi, Head of Department (Information Technology), Thakur College of Engineering and Technology, Mumbai and all the faculty members for conducting the project and their encouragement and co-operation has been a source of great inspiration. Also, I would like to thanks Thakur College of Engineering and Technology for providing me all the needful.

REFERENCES