Object Tracking using Statistic-based Feature Fusion Technique

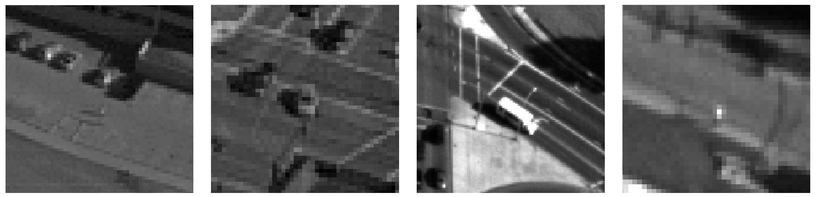
Evan Krieger, Paheding Sidike , Advisor: Dr. Vijayan Asari
Department of Electrical and Computer Engineering, School of Engineering

Introduction

Goal
Automatically track objects in wide area motion imagery (WAMI).

Constraints/Challenges
Very low resolution, presence of noise, illumination variation, occlusions, complex object motion, and complex object shapes.

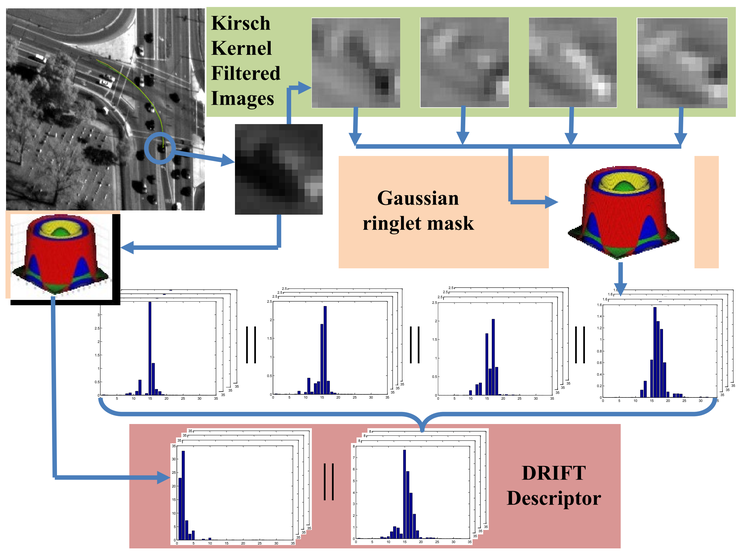
Proposed Innovation
To better combine features of the DRIFT algorithm through fusion based on past frame effectiveness



DRIFT Algorithm

Feature Extraction Method

1. Intensity image extracted from video frames
2. Four directional Kirsch kernels used to filter images
3. Gaussian ringlet masks used to create feature histograms for each Kirsch filtered image and the intensity image
4. Histograms from Kirsch filtered images concatenated to retain rotation invariance
5. All Histograms concatenated to create feature descriptor



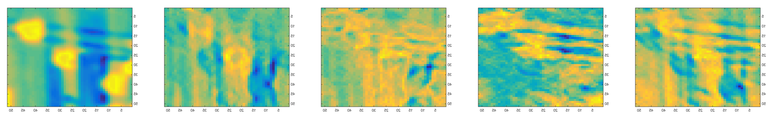
Fusion of Likelihood maps

- Each feature histogram is classified using Earth Mover's Distance
- A likelihood map is created for each feature for the search area
- Weights selected from previous frame performance
- Fusion based on variance ratio between target and background:

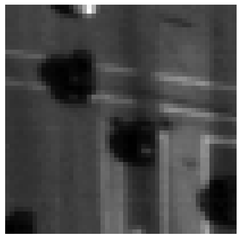
$$w_i = \frac{VR(L_i, p, q)}{\sum_{i=1}^{N} VR(L_i, p, q)}$$

$$VR(L_i, p, q) = \frac{var(L_i, (p+q)/2)}{[var(L_i, p) + var(L_i, q)]}$$

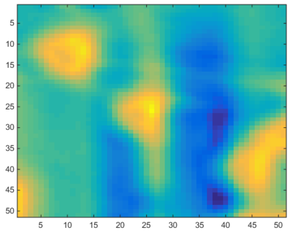
where object pixels, p, and background pixels, q



Feature Likelihood Maps



Search Area



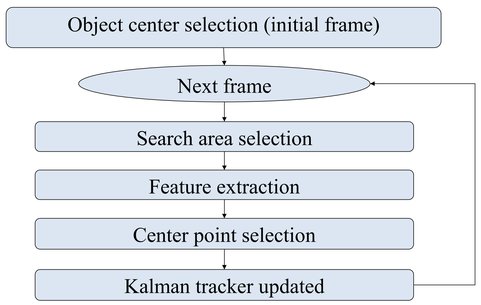
Fused Likelihood Map

Tracking process

$$\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$$

$$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

Kirsch Kernels



Kalman Tracker
Kalman tracker based on state equations of position and velocity to estimate position if an object is not detected

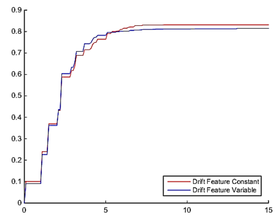
Each frame is processed by enhancement algorithm first.
Search area is selected based on Kalman Tracker, camera properties, and target properties.
Feature is extracted using DRIFT descriptor.
Classification is used to determine closest match.
The reference feature and Kalman Tracker are updated.

Results

- Results for constant weighting and variance ratio based weighting are compared
- 8 WAMI sequences were used for evaluations

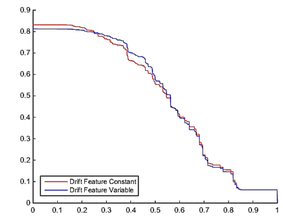
Spatial Robustness – Evaluated using different initial bounds

Precision Plot



Center Error Threshold

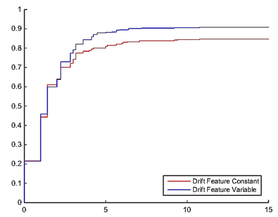
Success Plot



Overlap Threshold

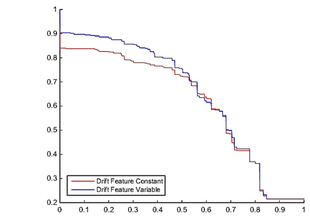
Temporal Robustness – Evaluated using different starting frames

Precision Plot



Center Error Threshold

Success Plot



Overlap Threshold

	Center Location Error	
	DRIFT Constant	Drift variable
CLIF1	11.22	**1.31**
CLIF2	**1.55**	1.86
CLIF3	**0.43**	0.49
CLIF4	1.76	**1.16**
BD1	**40.79**	42.66
BD2	0.70	**0.63**
BD3	**1.68**	1.96
BD4	25.96	**12.71**
Average	10.51	**7.85**

	Frame Detection Accuracy	
	DRIFT Constant	Drift variable
CLIF1	0.66	**0.74**
CLIF2	**0.72**	0.68
CLIF3	**0.88**	0.86
CLIF4	0.71	**0.78**
BD1	**0.36**	0.36
BD2	0.83	**0.85**
BD3	**0.72**	0.69
BD4	0.58	**0.70**
Average	0.68	**0.71**

UNIVERSITY of DAYTON *Vision Lab*
Kettering Laboratories
300 College Park
Dayton, OH 45469-0232

Center of Excellence
Computer Vision and Wide Area Surveillance Research

UD VISION LAB
www.visionlab.udayton.edu