Moving Object Removal in Unlabeled Image Databases

Victor Chen, Daniel Mendoza, Yechao Zhang

vchen36@stanford.edu, dmendo@stanford.edu, yechaoz@stanford.edu
Stanford University

Stanford
CS 229

Problem Statement

Moving object removal is a frequently experienced problem for image processing. This project aims to develop techniques for removing moving objects in a sequence of images with the same frame of reference. Currently, most conventional approaches use metrics such as mean, median, and frequency of pixel colors to extract background. This project applies modern machine learning techniques for moving object removal in unlabeled datasets where ground truth background is unknown.

One common approach is the Median Stack Filter (MSF) which inputs a sequence of images and outputs the median RGB value of each pixel

Left: Example image with moving objects highlighted; Right: Result of Median Stack Filter given 40 images

Approach

We implement unsupervised and semi-supervised techniques for background extraction.

Unsupervised Models:

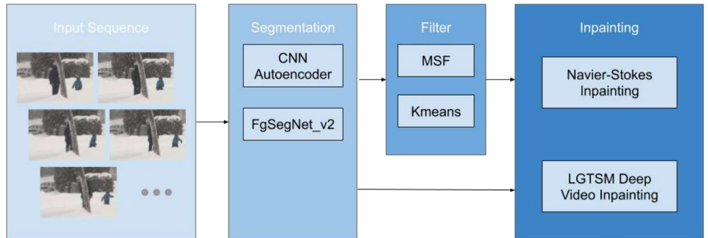
- KMeans Most Popular (MP) Cluster / Lowest Variance (LV) Cluster
- KMeans LV Inpainted Denoising

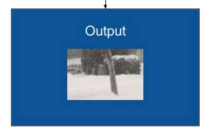
We implement a pipeline that utilizes both supervised and unsupervised methods where location of moving objects is inferred.

Models in Pipeline:

- Foreground segmentation -
  - CNN Autoencoder (deep learning)
  - Foreground Segmentation Network V2 (deep learning) [1]
- Filter -
  - Median Stack Filter
  - KMeans
- Inpainting -
  - Navier-Stokes
  - Learnable Gated Temporal Shift Module (deep learning) [2]

We evaluated our unsupervised approach and different combinations of the pipeline on three different test image sequences of 40 images each from the CDNET 2014 dataset containing 11 different video categories of ~70,000 frames [3].

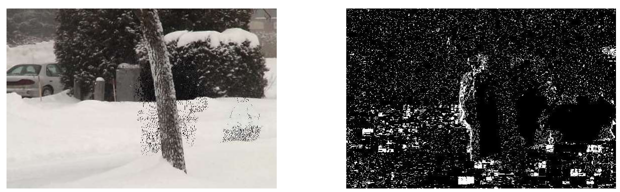




Pipeline for detecting the foreground, removing it, filtering the images, and applying inpainting to fill in the holes

Results

Left: KMeans MP Cluster output; Right: KMeans LV Cluster output



Left: KMeans LV Inpainted Denoising output; Right: Inpainting mask applied to KMeans LV Cluster output for noise reduction

Left: Pipeline output with CNN Autoencoder; Right: Pipeline output with FgSegNet_v2. Both images show output of pipeline using MSF and **Navier-Stokes Inpainting**.

Left: Pipeline output with CNN Autoencoder; Right: Pipeline output with FgSegNet_v2. Both images show output of pipeline using **LGTSM Video Inpainting**

Our pipeline allows interchanging of algorithms/techniques for each component, which can affect its performance. Removing the moving objects effectively relies upon an accurate foreground detector.

Conclusion

Given a sequence of images with moving objects, we implemented and evaluated unsupervised techniques as well as developed a pipeline algorithm with a foreground segmentation stage, filtering stage, and inpainting stage for background extraction.

- KMeans MP Cluster performs same as MSF.
- Kmeans LV Inpainted Denoising Cluster is able to remove most of the moving objects, but noise remains in the output.
- The components of our pipeline algorithm can be swapped out for different algorithms, and depending on the techniques used, we produce reasonable outputs.
- Knowing regions of foreground can produce more visually appealing results via pipeline algorithm

Future Work

In the future, we would look into different techniques for segmentation and inpainting, as the performance of these components determines the visual appeal of our results. Other configurations of our pipeline, like reversing the order of steps, could also be tested.

References

[1] Lim, Long Ang, and Hacer Yalim Keles. “Learning Multi-Scale Features for Foreground Segmentation.” Pattern Analysis and Applications (2019): n. pag. Crossref. Web.
[2] Ya-Liang Chang, Zhe Yu Liu, Kuan-Ying Lee, and Winston Hsu, “Learnable Gated Temporal Shift Module for Deep Video Inpainting”, in BMVC 2019.
[3] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, changedetection.net: A new change detection benchmark dataset, in Proc. IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, Providence, RI, 16-21 Jun., 2012.