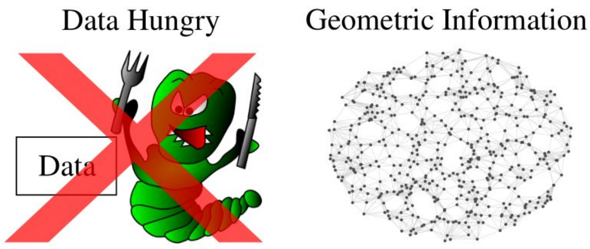
GraphBGS: Background Subtraction via Recovery of Graph Signals

Jhony H Giraldo and Thierry Bouwmans

International Conference on Pattern Recognition
10-15 January - Milan, Italy

Motivation and Problem

Motivation



Capturing geometrical information with graphs

- Deep learning approaches require large amounts of data to avoid overfitting.
- Graphs offer tools to exploit the geometrical structure of data.
- Most deep learning techniques in background subtraction do not evaluate their performance on unseen videos.

Motivation and Problem

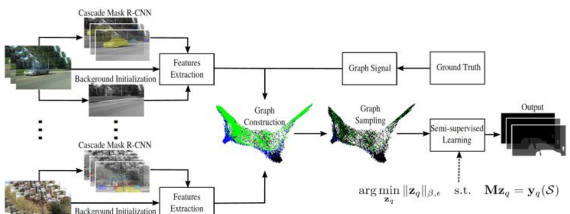
Problem

Background subtraction, moving object detection

- Background subtraction of unseen videos using semi-supervised learning techniques inspired from graph signal processing.

Algorithm

Overview

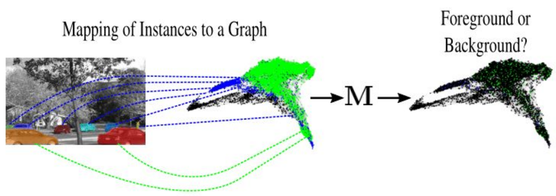


Pipeline of our algorithm

- Our proposed algorithm consists of several components including:
  - Instance segmentation using Cascade Mask R-CNN [Cai and Vasconcelos (2019)]
  - Background initialization using a temporal median filter [Piccardi (2004)]
  - Feature extraction with motion [Lucas et al. (1981)], texture [Ojala et al. (2002)], and intensity features
  - Semi-supervised learning inspired from Sobolev norm minimization.

Algorithm

Instances and Graph

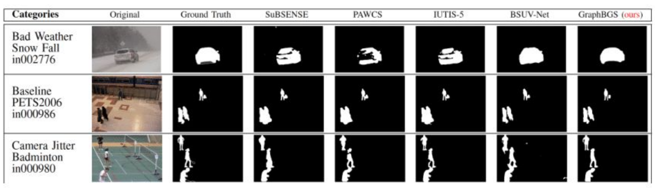


Cascade Mask R-CNN, nodes representation, sampling.

- Instance masks are mapped as nodes of a graph G.
- The representation of the nodes is given by:
  - Motion features
  - Texture features
  - Intensity features
  - Background features.
- The graph is constructed with a k-nearest neighboors with k= 30.
- The graph signal associated to the problem of background subtraction is sampled for a evaluation with unseen videos.

Experimental Framework

Results I



Some visual results on CDNet2014.

Challenge	FTSG	SuBSENSE	PAWCS	WisenetMD	IUTIS-5	SemanticBGS	FgSegNet v2	BSUV-Net	GraphBGS
Bad Weather	0.8228	0.8619	0.8152	0.8616	0.8248	0.8260	0.7952	**0.8713**	**0.9085**
Baseline	0.9330	0.9503	0.9397	0.9487	0.9567	**0.9604**	0.6926	**0.9693**	0.9535
Camera Jitter	0.7513	0.8152	0.8137	0.8228	0.8332	**0.8388**	0.4266	0.7743	**0.8826**
Dynamic-B	0.8792	0.8177	**0.8938**	0.8376	0.8902	**0.9489**	0.3634	0.7967	0.8353
I-O Motion	**0.7891**	0.6569	0.7764	0.7264	0.7296	**0.7878**	0.2002	0.7499	0.5036
Low-F rate	0.6259	0.6445	0.6588	0.6404	**0.7743**	**0.7888**	0.2482	0.6797	0.6022
PTZ	0.3241	0.3476	0.4615	0.3367	0.4282	0.5673	0.3503	**0.6282**	**0.7993**
Shadow	0.8832	0.8986	0.8913	0.8984	0.9084	**0.9478**	0.5295	0.9233	**0.9712**
Thermal	0.7768	0.8171	0.8324	0.8152	0.8303	0.8219	0.6038	**0.8581**	**0.8594**
Overall	0.7539	0.7566	0.7870	0.7653	0.7923	**0.8320**	0.4158	0.8056	**0.8128**

Comparisons of average F-measure in CDNet2014. The best and second best performing method for each challenge are shown in red and blue, respectively.

Summary and Future Work

- This work introduces concepts of graph signal processing in the problem of background subtraction.
- Our algorithm is composed of: a Cascade Mask R-CNN for instance segmentation, a temporal median filter for background initialization, feature extraction for node representation, a k-nearest neighbors for graph construction, and minimization of the Sobolev norm of graph signals for the semi-supervised learning algorithm.
- For future work:
  - A generalized theory of graph signal processing [Ji and Tay (2018)] can be used to extend the graph signals y to fuzzy concepts [El Baf et al. (2008)], leading to a richer representation of moving and static objects.
  - Another important direction is to study an inductive framework [Hamilton et al. (2017)] for GraphBGS, trying to get a real-time implementation of GraphBGS [Cioppa et al. (2020)].
  - Perhaps, concepts of graph signal processing such as active semi-supervised learning [Anis et al. (2018)] could lead in new developments in the field of computer vision and end-to-end architectures for video analysis with semi-supervised learning.