

Content Aware Video Summarization

Michael N. Lombardo
Michael.Lombardo@uoit.net

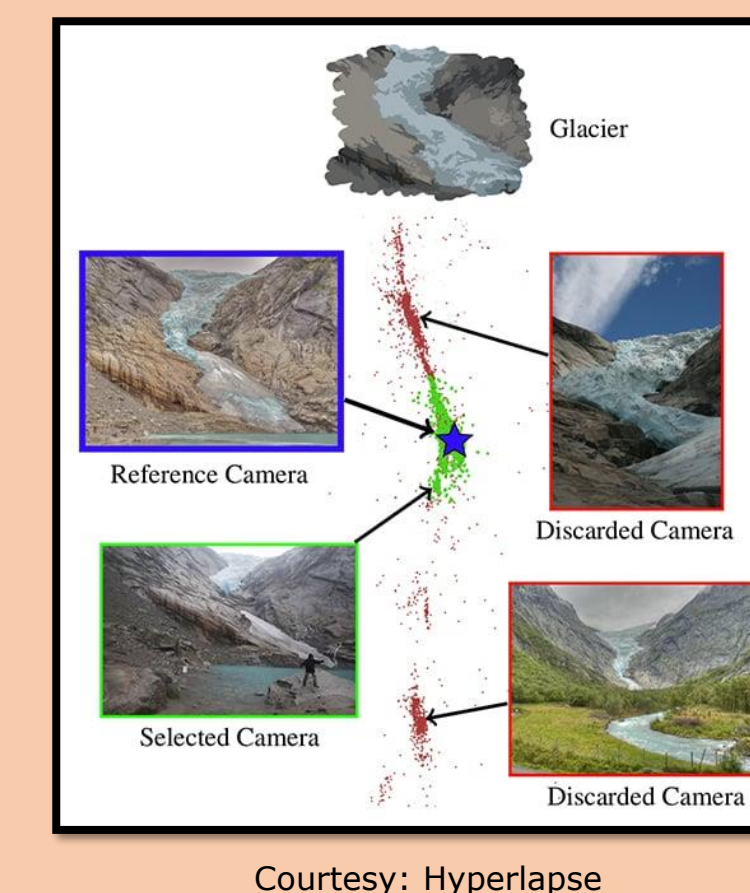
Faculty of Science
University of Ontario Institute of
Technology
Supervisor: Faisal Qureshi

Introduction

- Video recording devices are now pervasive, and as a consequence, our ability to record our lives in high-definition has grown simple.
- Raw video footage however, is rarely ready for general viewing.
 - We urgently need theory, methods and tools for automatic video editing or summarization.
- This work aims to develop new methods for Content Aware Video Summarization by leveraging two recent technologies:
 - Hyperlapse for *fast-forwarding* video playback
 - YOLO9000 for object detection.

Hyperlapse

- Fast-forwarding video play back often amplifies camera movement, creating a dizzying viewing experience.
- Hyperlapse exploits structure-from-motion techniques to:
 - construct 3D scene structure
 - infer camera movement.
- The video is re-rendered so as to achieve a desired frame rate along a hallucinated camera path that minimizes jitter.



Acknowledgements

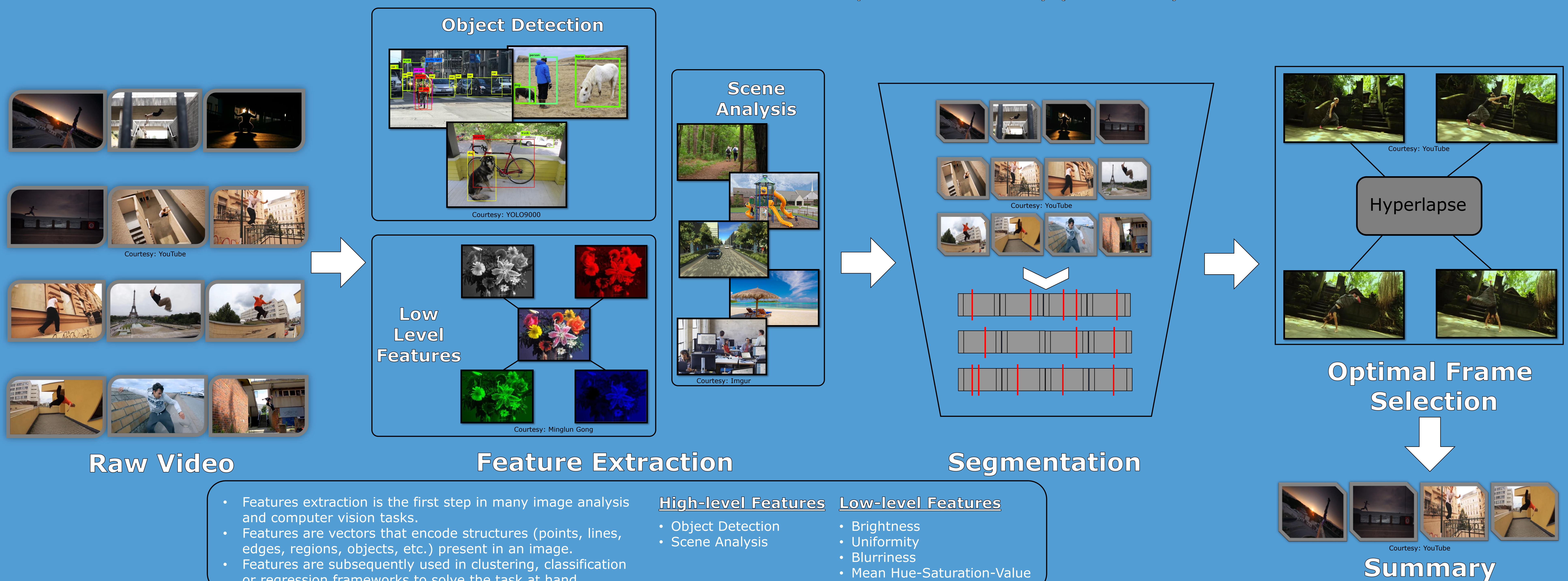
A special thank you to the Natural Sciences and Engineering Research Council and the Visual Computing Lab for funding this research.

Future Work

- Action recognition to improve segment quality
- Caption guided summarization
- Sports video dataset for summarization benchmarks

Video Summarization Framework

Our task: decided whether or not a frame should be kept in the video summary. (Classification)



Yolo9000

- Fast, accurate general purpose object detector.
- You Only Look Once (YOLO9000) recognizes more than 9000 object categories by jointly optimizing detection and classification.
- Deep learning.

