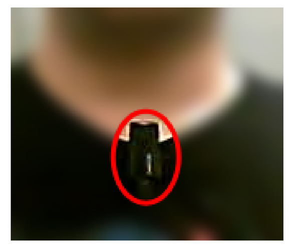
Novelty Detection from an Ego-Centric Perspective

Omid Aghazadeh   Josephine Sullivan   Stefan Carlsson
Computer Vision and Active Perception laboratory, Stockholm, Sweden

Introduction

- What is worth remembering in a data-set of everyday life videos? (repeated activities)

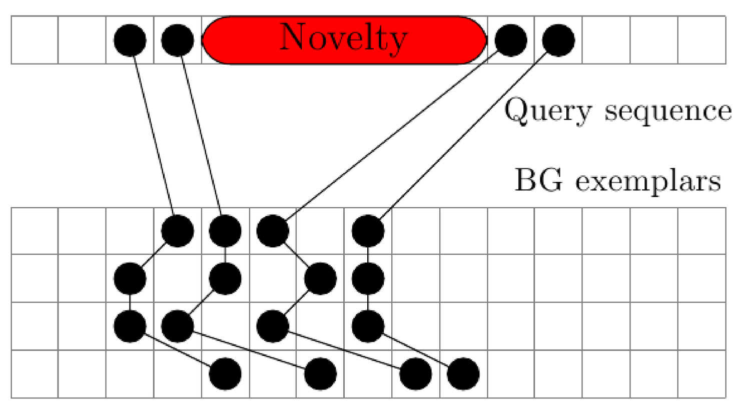


each row shows samples from a sequence of a subject walking to work
each sequence corresponds to a different day

- Events can be divided to repeated(background) and unique(novel)
- Background can be compressed: represented by a few exemplars
- Novel events need to be memorized in addition to the bg exemplars
- Novelty detection can be used as a memory selection process

Novelty Detection via Sequence Alignment

- Novel frames do not have correspondences in bg exemplars (reference sequences)
- Establishing correspondences between frames of videos determines novelties



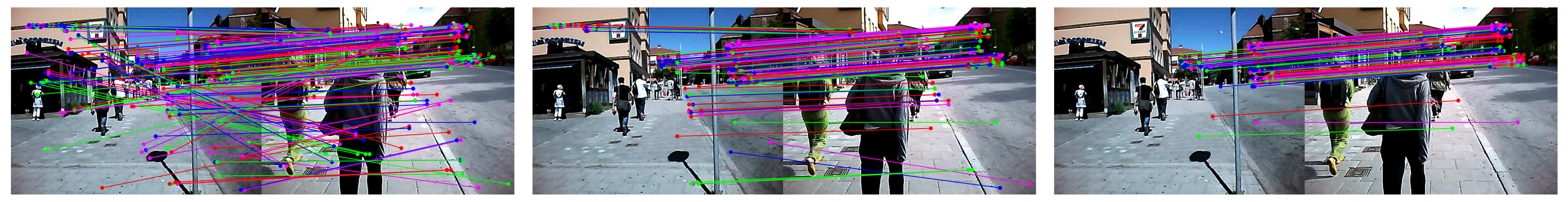
- Aligning sequences establishes correspondences
- A pairwise similarity measure defines a similarity matrix
- Sequences are aligned using **Dynamic Time Warping** on the similarity matrix

Dataset

- 31 videos of on average 5 minutes of a subject walking to work
- Each frame is manually labeled with a virtual location
- 4 sequences were manually identified to contain novelties
- Significant illumination/viewpoint variations
- Non-static environment

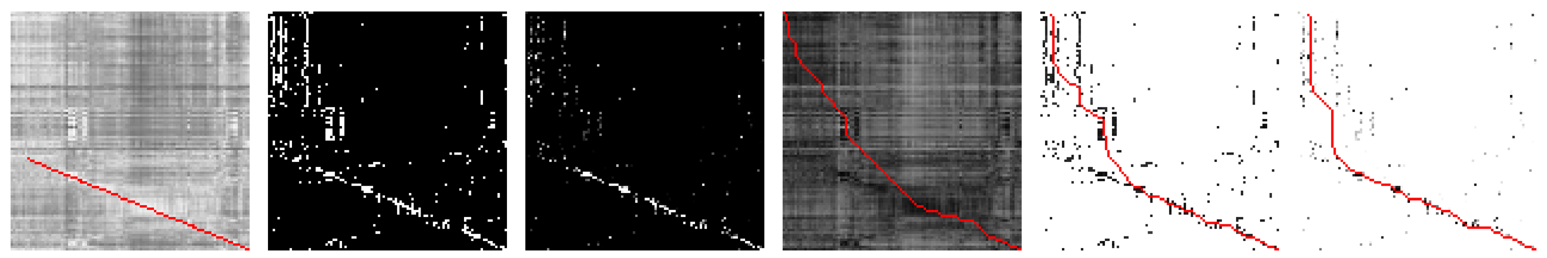
Similarity Measure for novel Ego-Motion detection

- Similarity in ego motion ↔ Similarity in view point
- Kernel operating on Vector Space model (*V.S.*) of images
  - Inexpensive but not accurate enough
- **Geometric similarity** (*G.S.*) between two images
  - Robust estimation (Prosac) + Epipolar geometry (5 point)
  - Expensive but accurate



250 putative matches   inliers wrt H   inliers wrt H and E

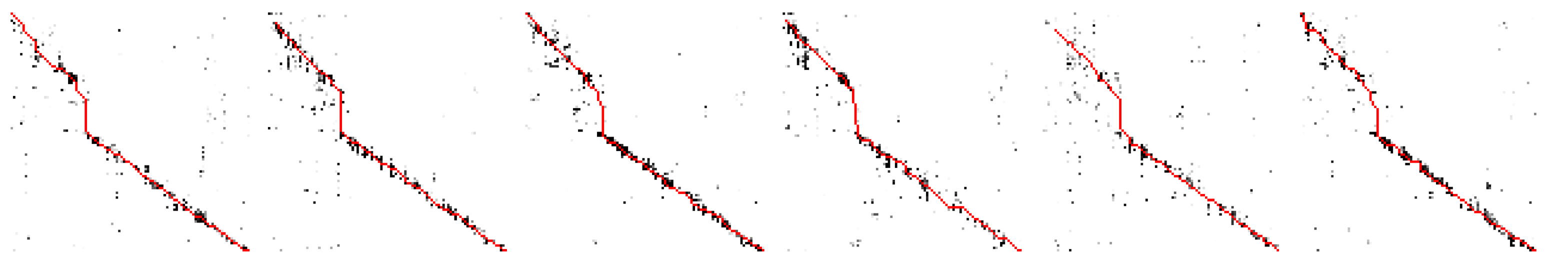
- A loose Homography plays a *depth regularization* role for Epipolar constaint
- Sparse similarity matrix: evaluate it on V.S.-based KNNs of each frame



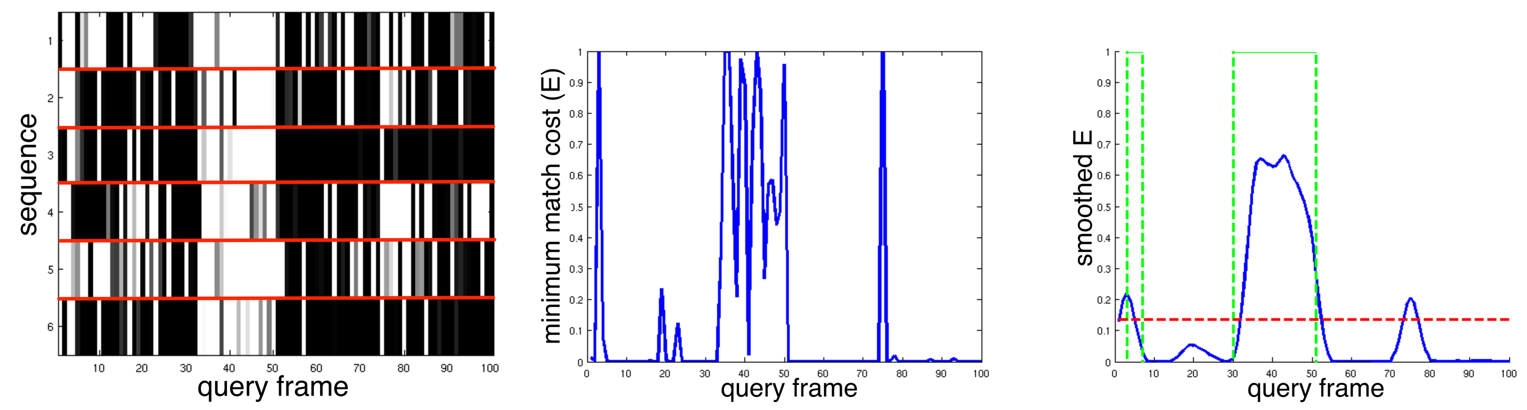
dense V.S., sparse V.S., sparse G.S. and the resulting alignments
vertical axis: query frames, horizontal axis: reference frames

Temporally consistent novelty detection

- After alignment, each frame will be associated with a matching cost
- Stacking them in a matrix → match cost
- Getting the minimum → minimum match cost
- **Thresholding the smoothed minimum match cost** → novelties



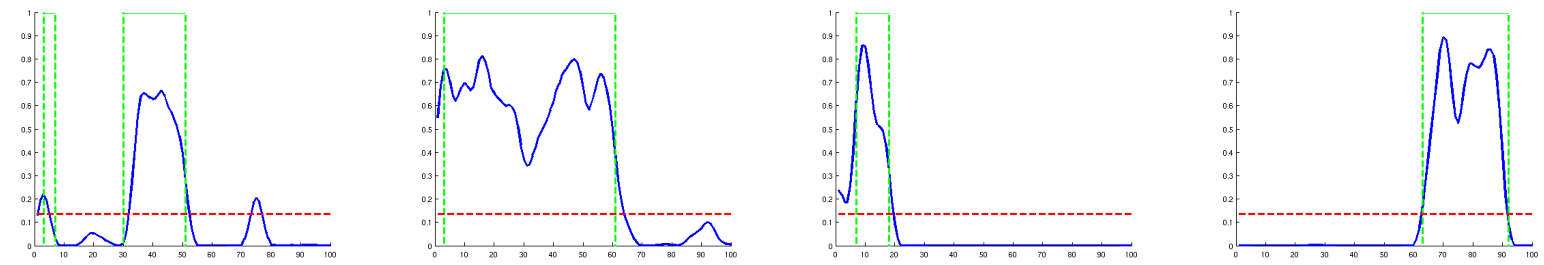
aligning reference sequences with a query sequence



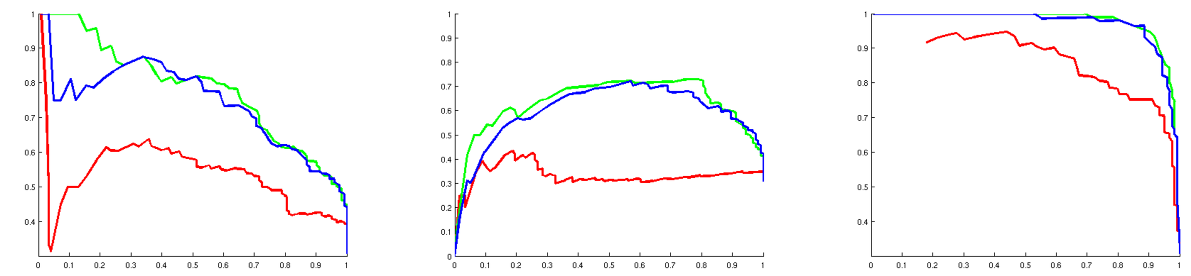
match cost, minimum match cost and smoothed minimum match cost

Evaluation of G.S. and novelty detection

- Evaluation on the 4 sequences which contained novelty (400 frames)



ground truth, smoothed minimum match cost, a constant threshold



dense V.S.   sparse V.S.   sparse G.S.

- PR curves using 1, 6 and 10 reference sequences

Qualitative Results

(meeting a friend)

(ice cream shopping)

Conclusions

- Novelty detection can be used as a memory selection process
- Similarity measures define similarity matrices
  - Similarity in viewpoint ↔ Similarity in ego-motion
- Novel events in videos can be detected by sequence alignment based on a similarity measure

This work was supported by The Swedish Foundation for Strategic Research in the project Wearable Visual Information Systems.
{omida,sullivan,stefanc}@csc.kth.se