Temporal Causality for Analysis of Visual Events

Karthir Prabhakar, Sangmin Oh, Ping Wang, Gregory D. Abowd, James M. Rehg
Georgia Institute of Technology

presented by
Shashwat Purushottam
Outline

- Introduction
- Motivation
- Some background & context
- Key Idea
- Algorithm Procedure
- Point processes - Visual words
- Temporal Causal Analysis
- Results
- Conclusion
Introduction

● What is causality?
● Analysis of motion sequences to establish meaningful relationships between "visual events", like the associated temporal dependencies.
● Example - The game of soccer.
● Existence of domain models to explain causality and analyse events otherwise
● Granger Causality
● In this paper we present a novel, data-driven analysis of causality in video.
Motivation

- Establishing causality at the event level of granularity is extremely useful in applications like Video Segmentation.
- Computational complexity is extremely low.
- It is able to aggregate information across a long time interval; sustains continual analysis.
- Quasi periodic and periodic or repetitive motion can be captured very well.
- "Video events" like social games can incur the benefits of improved retrieval and classification performance.
Some background & context

- Pixel based representation - any good?  
  - Expensive.
- The pattern of general videos and their analysis makes it conducive to work with an event based representation.
- Space-time visual words are used to represent the motion video sequence by recognizing low-level visual events - Similar in object categorization.
- Temporal causal analysis to segment visual events into independent groups.
B & C Continued

- No prior work on causal analysis of such sequences using space-time visual words based on temporal co-occurrences.
- Some prior work on causal analysis - baseball action video - domain knowledge - 1:1 hand tracking - human annotations
- Needs to work with general video content in an unsupervised fashion such that grouping depends mainly on temporal interaction as opposed to user interactions.
- Granger's causality theorem
Key Ideas

3 main contributions in the paper:

- a novel representation of video content which encodes the dynamics of visual events as visual words which can be instantiated as a multivariate point process.
- nonparametric pairwise causality test for point-processes can be used to partition visual events into independent groups.
- Demonstrate both qualitative segmentation performance and quantitative improvements in retrieval and categorization of social games.
Algorithm Procedure

It is divided into two steps:

- **Point-process representation of video** -
  -- Each visual word occurs in a subset of frames and can therefore be represented as a point process.
  -- The number of occurrences of an event 'i' in a time interval (0,t] is the point process function of the visual word that represents that event.
  -- m visual words in a sequence form an m dimensional multivariate point process.

- **Spectral representation of point processes** -
  -- The statistical relationship between a pair of point processes can be captured by its cross-spectral density function which can be estimated from non parametric sample data.
Point processes - visual words
Temporal Causal Analysis

(a) Spectral Matrix  
(b) Causal Measures  
(c) Thresholding  

(d) Causal Matrix  
(e) Causal Graph

Figure 2. Visualization of Temporal Causal Analysis.
Results

(a) Ball-throw sequence with other noise motion

(b) Handshake sequence from the movie Forrest Gump
Conclusion

- The ability to characterize the temporal structure of video events in an unsupervised manner is key to video analysis.
- It can ameliorate the need for training data.
- The authors have shown that a nonparametric formulation of Granger causality can be used to identify patterns of interaction between repeating events, and partition them into independent causal sets.
- Improved performance in retrieving and classifying social games from a video.