Change Point Detection for Poisson Time Series Images with Applications to Astronomy and Astrophysics

Thomas C. M. Lee

Department of Statistics
University of California, Davis

Joint work with Vinay Kashyap and Cong Xu

November 27, 2018

Outline

Outline

Problem Description

- Abrupt changes of intensities, energy spectra and spatial patterns in astronomical sources are of interest to astrophysics.
- E.g., Chandra observations of the Orion Nebula cluster.
- Goal: detect change points in the time direction.

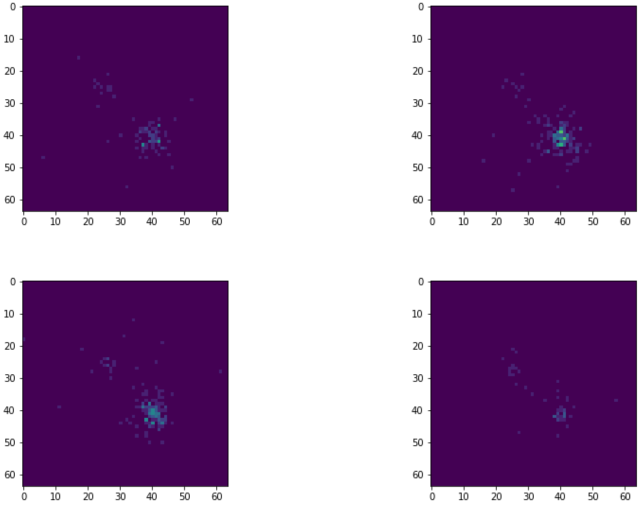
Data

- Data are usually obtained in the form of a list of photons, with four attributes (x, y, t, w):
  - the 2D spatial coordinates (x, y) where the photons were detected
  - the times t they were recorded
  - wavelengths w (i.e., energies)
- We first focus on the spatial and the time attributes of the photons.
- The case for spectral and time has been done - AutoMARK.

Binning

- We first bin the original data into a 3D rectangular grid of equi-volumned boxes (i.e., voxels)
- Thus a 3D table of photon counts indexed by (x, y) and t.
- Can be viewed as a time series of images with photon counts as pixel values.

Binned Data: at four time points



Assumptions: Poisson

- Emission of photons follows a non-homogeneous Poisson process.
- So Poisson counts are independent.
- Observed images $\{y_{i,t}\}, i = 1, ..., n, t = 1, ..., T$ satisfy

$$y_{i,t} \sim \text{Poisson}(g_{i,t}),$$

  where
  - n is the number of pixels in one image,
  - T is number of images,
  - $g_{i,t}$: unknown true value at location i and time t.
  - (i is bivariate: x and y)

Assumptions: Homogeneous Periods

- The T images can be partitioned into $(K + 1)$ homogeneous periods by K change points $\tau_1, \tau_2, ..., \tau_K$.
- (Let $\tau_0 = 0, \tau_{K+1} = T$.)
- Images from the same period are assumed to have the same unknown $g_{i,t}$.

Assumptions: Piecewise Constant for $g_{i,t}$

- Model g with a 2D piecewise constant function:

$$g_{i,t} = \sum_{k=1}^{K+1} I_{\{t \in (\tau_{k-1}, \tau_k]\}} \sum_{\gamma=1}^{m^{(k)}} f_\gamma^{(k)} I_{\{i \in R_\gamma^{(k)}\}}, \quad \forall i, t.$$

- $f_\gamma^{(k)}$: unknown Poisson parameter for the γth region of the kth period.
- $R_\gamma^{(k)}$: index set for the γth region in the kth period.
- Note: $R_\gamma^{(k)} \subseteq \{1, ..., n\}$.
- $m^{(k)}$: number of regions in the image specified by $R_\gamma^{(k)}$

As a Model Selection Problem

- Given data $\{y_{i,t}\}$ such that

$$y_{i,t} \sim \text{Poisson}(g_{i,t}) \quad \forall i, t,$$

  we want to obtain an estimate of $g_{i,t}$.
- In other words, we want to estimate
  - number K and locations of change points τ_k's
  - segmentations (object boundaries) of the images $R_\gamma^{(k)}$
  - Poisson parameters $f_\gamma^{(k)}$.
- The estimation is a model selection problem.
- We will use the minimum description length (MDL) principle.

Outline

What is MDL?

- MDL defines the best model as the one that produces the best compression (minimization of the code length) of the data.
- Code length (or description length): amount of hardware memory to store an object

$$\begin{aligned} \text{CL(data)} &= \text{CL(fitted model)} + \text{CL(data given fitted model)} \\ &= \text{CL(fitted model)} + \text{CL(residuals)} \\ &= \text{CL(explained by model)} + \text{CL(not explained by model)} \end{aligned}$$

  - First term: model complexity
  - Second term: data fidelity (essentially log-likelihood)
- The best fitting model is defined as the minimizer of CL(data).
- Why does MDL work as a model selection method?

MDL for Homogeneous Poisson Image Series

- That is, no change point and all T images share the same true 2D piecewise constant function.
- Our model reduces to, $\forall i$ and t:

$$
\begin{aligned}
y_{i,t} &\sim \text{Poisson}(g_{i,t}) \\
g_{i,t} &= \sum_{\gamma=1}^{m} f_\gamma I_{\{i \in R_\gamma\}}
\end{aligned}
$$

MDL for Homogeneous Poisson Image Series

- m: number of regions
- $R = \{R_\gamma | \gamma = 1, ..., m\}$ as the partition (i.e., object boundaries)

$$\mathrm{MDL}(m, R) = m \log(n) + \frac{\log(3)}{2} \sum_{\gamma=1}^{m} b_\gamma + \frac{1}{2} \sum_{\gamma=1}^{m} \log(T a_\gamma) - \sum_{\gamma=1}^{m} \sum_{t=1}^{T} \sum_{i \in R_\gamma} y_{i,t} \log(\hat{f}_\gamma^{(T)})$$

- a_γ: the "area" (number of pixels) of region R_γ
- b_γ: the "perimeter" (number of pixel edges)
- $\hat{f}_\gamma^{(T)}$: the sample mean of region R_γ when (m, R) are given, which is

$$\hat{f}_\gamma^{(T)} = \frac{1}{T a_\gamma} \sum_{t=1}^{T} \sum_{i \in R_\gamma} y_{i,t}$$

General Case: Change Point Detection

- Unknown K change points $\tau_1, \tau_2, ..., \tau_K$. Let $\tau_0 = 0, \tau_{K+1} = T$.

$$
\begin{aligned}
y_{i,t} &\sim \text{Poisson}(g_{i,t}) \\
g_{i,t} &= \sum_{k=1}^{K+1} I_{\{t \in (\tau_{k-1}, \tau_k]\}} \sum_{\gamma=1}^{m^{(k)}} f_\gamma^{(k)} I_{\{i \in R_\gamma^{(k)}\}}
\end{aligned}
$$

- Let $\lambda_j = \tau_j / T$ for $j = 0, 1, ..., K+1$.
- $\lambda = (\lambda_0, ..., \lambda_{K+1})$.
- Numbers of regions within each period: $m = (m^{(1)}, ..., m^{(K+1)})$.
- Object boundaries for all periods: $R = (R^{(1)}, ..., R^{(K+1)})$, where $R^{(k)} = \{R_\gamma^{(k)} | \gamma = 1, ..., m^{(k)}\}$ for $k = 1, ..., K+1$.

General case: Change Point Detection

The overall MDL is

$$\mathrm{MDL}_{\mathrm{overall}}(K, \lambda, m, R) =$$
$$K \log(T) + \sum_{k=1}^{K+1} \mathrm{MDL}(m^{(k)}, R^{(k)}, \lambda_{k-1}, \lambda_k, \hat{f}(R^{(k)}, \lambda_{k-1}, \lambda_k)),$$

where $\mathrm{MDL}(m^{(k)}, R^{(k)}, \lambda_{k-1}, \lambda_k, \hat{f}(R^{(k)}, \lambda_{k-1}, \lambda_k))$ is the homogeneous MDL for images from $(T\lambda_{k-1} + 1)$ to $T\lambda_k$.

Outline

Consistency in Homogeneous Case

Theorem (Homogeneous case; i.e., no change point)

As $T \to \infty$,

$$\hat{m} \to m^0 \quad and \quad \hat{R} \to R^0,$$

where m^0 *and* R^0 *are the corresponding true values.*

Consistency in General Case

Theorem (General case with change points)

As $T \to \infty$

$$\hat{K} \to K^0, \quad \hat{\lambda}_k \to \lambda_k^0 \quad a.\ s., \quad \hat{m}^{(k)} \to m^{0(k)} \quad and \quad \hat{R}^{(k)} \to R^{0(k)},$$

where K^0, λ_k^0, $m^{0(k)}$ *and* $R^{0(k)}$ *are the corresponding true values.*

To the best of our knowledge, this is the first consistency result for simultaneous change point detection and segmentation for time series of images.

Outline

Optimization: Very Challenging

- Global minimization of $\text{MDL}_{\text{overall}}(K, \lambda, m, R)$ is practically infeasible for not-so-small T and n.
- Two stages:
  1. change points detection
  2. image segmentation
- techniques involved:
  - seeded region growing
  - region merging
  - forward selection/backward elimination
  - some heuristics, and others ...
- still a bit slow, needs improvement (later)

Outline

Simulation

To make the setup more realistic:

- Fit a model to a real dataset with the algorithm described before.
- Generate datasets based on the fitted model.
- Apply the algorithm to the generated datasets

Simulation

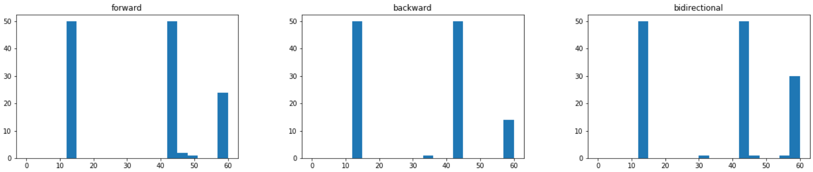
Real data:

- ObsID: 04373_1
- $n = 64 \times 64$
- $T = 60$

Fitted model:

- Change points: 12, 42, 57

Simulation: results from 3 algorithms



Simulation: further results

Method	% correct $\hat{K}$	% exactly the same
Forward	54%	48%
Backward	30%	28%
Bidirectional	58%	56%

Outline

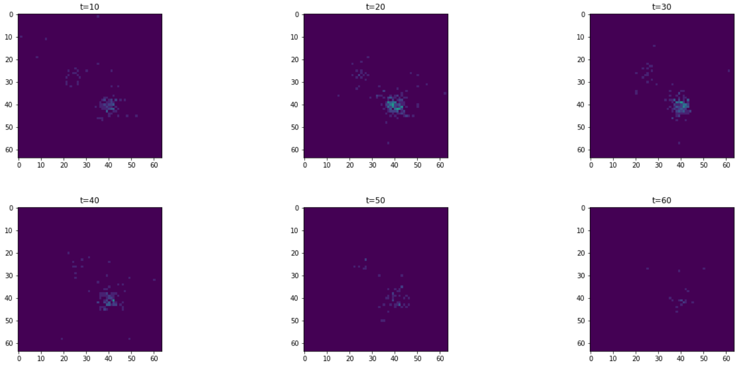
1. Background and Modeling
2. The Proposed Method
3. Theoretical Properties
4. Optimization Algorithm
5. **Practical Performance**
6. Concluding Remarks

Application to real data

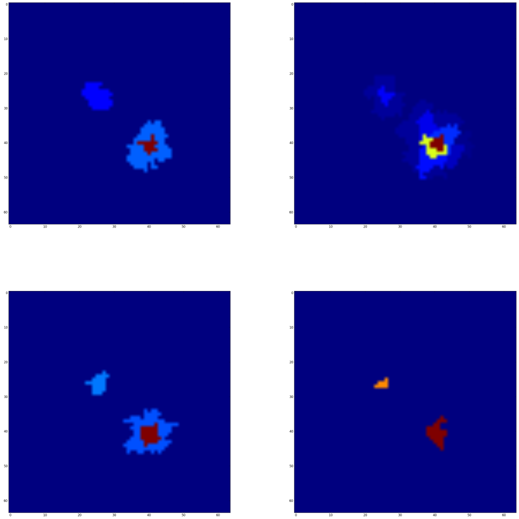
Real data:

- ObsID: 04373_1
- $n = 64 \times 64$
- $T = 60$

Application to real data



- Change points: 12, 42, 57



Outline

Concluding Remarks

Past:

- developed a method for simultaneous change point detection and segmentation for time series of images
- to the best of our knowledge, first time consistency results are established

Present:

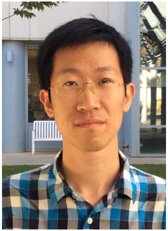
- extension to exponential family (done)
- incorporate spectral information; i.e., include w which leads to a 4D version of the problem (done)
- fused lasso to speed up the algorithm (working on it)

Future:

- relax the piecewise constant assumption; i.e., piecewise polynomial
- successful astro-X applications!

Thank You!

- Thank you!
- collaborators:



Cong Xu

Vinay Kashyap

you?