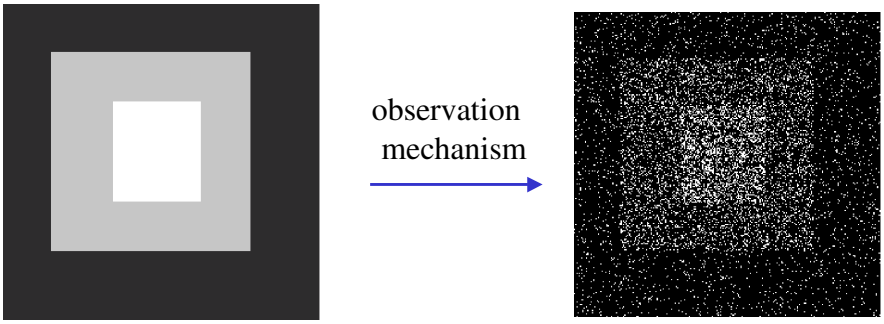
1

Recursive Segmentation of Poisson Fields

Mário A. T. Figueiredo
"Instituto Superior Técnico"
Lisboa
PORTUGAL

2

Poisson image segmentation: an inverse problem



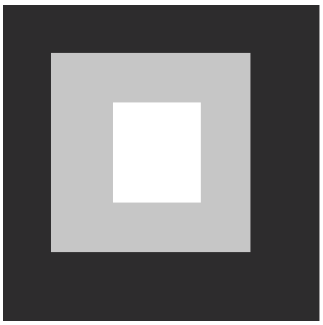
f, intensity
(piece-wise constant)

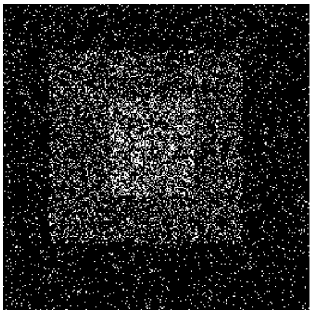
g, observed counts

Inverse problem

Goal: given **g**, find **f**

3 The model selection issue





f, intensity

g, observed counts

Inverse problem

Model selection question: how many (constant intensity) regions ?

4 The minimum description length (MDL) criterion

Introduction to MDL

g observed data

encoder

compressed data

decoder

g

Rationale:

short code $\Leftrightarrow$ good model

long code $\Leftrightarrow$ bad model

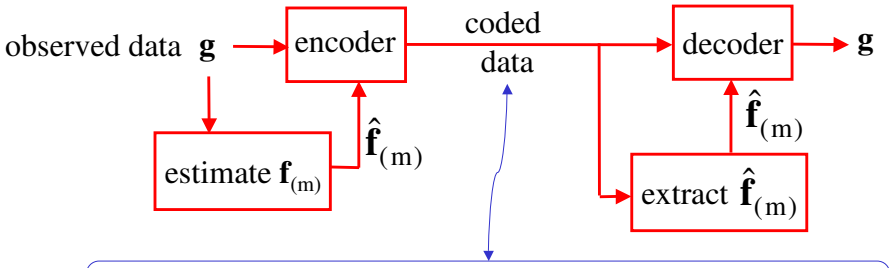
code length $\Leftrightarrow$ model adequacy

Several flavors: Rissanen 78, JRSS 87
Rissanen 96,
Wallace and Freeman (MML), JRSS 87

5 Formalizing the MDL criterion

Competing models: $\{p(\mathbf{g}\,|\,\mathbf{f}_{(m)}, m)\,, m = m_1, m_2, ..., m_K\}$

Assumption: given $\mathbf{f}_{(m)}$, both encoder and decoder know how to build the same code.



coded data = code(m) + code($\hat{\mathbf{f}}_{(m)}$ | m) + code(g |$\hat{\mathbf{f}}_{(m)}$, m)

MDL principle:
choose m and $\hat{\mathbf{f}}_{(m)}$ so that length(coded data) is shortest

6 Formalizing the MDL criterion

coded data = code(m) + code($\hat{\mathbf{f}}_{(m)}$ |m) + code(g |$\hat{\mathbf{f}}_{(m)}$,m)

$$L(m, \mathbf{f}_{(m)}, \mathbf{g}) = \underbrace{L(m)}_{\text{Usually constant}} + L(\mathbf{f}_{(m)}\,|\,m) + \underbrace{L(\mathbf{g}\,|\,\mathbf{f}_{(m)}, m)}_{\text{Ideal code-length}}$$

$$-\log p(\mathbf{g}\,|\,\mathbf{f}_{(m)}, m)$$

MDL criterion

$$(\hat{m}, \hat{\mathbf{f}}_{(\hat{m})})_{\text{MDL}} = \arg \min_{m, \mathbf{f}_{(m)}} \{L(\mathbf{f}_{(m)}\,|\,m) - \log p(g\,|\,\mathbf{f}_{(m)}, m)\}$$

What about $L(\mathbf{f}_{(m)}|m)$?

Usually, this requires asymptotic arguments (e.g., the well-known (1/2) log n)

7 The incompleteness issue in MDL

Classical MDL (Rissanen, 1978)

$$(\hat{m},\hat{\mathbf{f}}_{(\hat{m})})_{MDL} = \arg \min_{m,\mathbf{f}_{(m)}} \{\frac{m}{2}\log(n) - \log p(\mathbf{g}\,|\,\mathbf{f}_{(m)})\}$$

Problems: not valid for small samples (asymptotic criterion);
assumes a redundant (called *incomplete*) code (Rissanen, 1996).

Incompleteness: in code-length $-\log p(\mathbf{g}\,|\,\mathbf{f}_{(m)})$

1. Transmitter estimates $\mathbf{f}_{(m)}$ from $\mathbf{g}$
2. Transmitter sends $\hat{\mathbf{f}}_{(m)}$ to receiver
3. Now, receiver knows that the only possible $\mathbf{g}$'s are those that can lead to that $\hat{\mathbf{f}}_{(m)}$
   Only needs to build code for those $\mathbf{g}$'s

8 The incompleteness issue in MDL

Example: $\mathbf{g} = \{g_1,g_2,\ldots,g_{10}\}$ $\quad g_i \in \{0,1\}$

10 i.i.d. observations of a Bernoulli source

$\text{Prob}(g_i=1) = f$

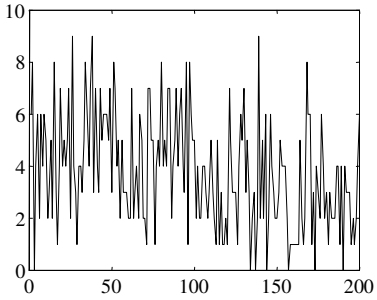
1. Transmitter computes $\hat{f} = \frac{1}{10}\sum_{i=1}^{10} g_i$
2. Transmitter sends $\hat{f}$ to receiver (rational number)
3. Not all $\mathbf{g}$'s need be coded.
   Only those that have $10\,\hat{f}$ ones: $L(\mathbf{g}) = \log\binom{10}{10\,\hat{f}}$

Concrete example: $\mathbf{g} = \{1,0,0,0,1,0,1,0,0,0\}$, 3 ones $\longrightarrow \hat{f} = \frac{3}{10}$

$\log_2\binom{10}{3} = 6.9$ bits versus $-\log_2 p(\mathbf{g}\,|\,\hat{f}) = -\log_2[0.3^3(1-0.3)^7] \approx 8.8$ bits

9 MDL for change-point detection in Poisson sequences

Basic question: given a sequence of independent Poisson counts



$$\mathbf{g} = \{g_1, \ldots, g_N\}$$

Is there a change in the intensity ?
If so, where ?

We have N models:

$$m_0 : \; p(\mathbf{g} \mid \lambda, m_0) = \prod_{i=1}^{N} \text{Poisson}(g_i \mid \lambda) \quad \text{(no change)}$$

$$m_1, \ldots, m_{N-1} : \; p(\mathbf{g} \mid \lambda_a, \lambda_b, m_k) = \prod_{i=1}^{k} \text{Poisson}(g_i \mid \lambda_a) \prod_{i=k+1}^{N} \text{Poisson}(g_i \mid \lambda_b)$$

10 MDL for change-point detection in Poisson sequences

Fact: Let $t = \sum g_i$ be the total count

Conditioned on t, the g_i's follow a multinomial distribution

$$p(g_1, \ldots, g_N \mid t) = \text{Multi}(g_1, \ldots, g_N \mid t; \alpha_1, \ldots, \alpha_N)$$

Approach: rather than the Poisson model,
condition on the total count, and use the multinomial

Remark: The total count is ancillary for the change detection problem

11 MDL for change-point detection in Poisson sequences

In multinomial terms the models are:

$$p(\mathbf{g} \mid t, m_0) = \text{Multi}(g_1,...,g_N \mid t, \frac{1}{N},...,\frac{1}{N})$$

$$p(\mathbf{g} \mid t, m_k) = \text{Multi}(g_1,...,g_N \mid t, \underbrace{\alpha_a,...\alpha_a}_{k}, \underbrace{\alpha_b,...,\alpha_b}_{N-k})$$

where $\alpha_a = \dfrac{\lambda_a}{k\lambda_a + (N-k)\lambda_b}$ $\qquad$ $\alpha_b = \dfrac{1-k\alpha_a}{N-k}$

Notice that $\lambda_a = \lambda_b \Rightarrow \alpha_a = \alpha_b = \dfrac{1}{N}$

12 MDL for change-point detection in Poisson sequences

Our MDL approach:

1. given $\mathbf{g} = \{g_1,...,g_N\}$ compute $t = \sum g_i$
2. transmit t (e.g., with an arbitrary technique, e.g. Elias code)
3. compute the description length under model m_0

$$L(\mathbf{g} \mid t, m_0) = -\log \binom{t}{g_1\ g_1 \cdots g_N} + t \log N$$

This is the total description length, because there is no parameter

4. compute the description length under models $m_1,\ldots,m_{N-1}$

13 MDL for change-point detection in Poisson sequences

The code-length, under model $m_k \in \{m_1,...,m_{N-1}\}$

1. Estimate $\alpha_a \longrightarrow \hat{\alpha}_a = \dfrac{\sum_{i=1}^{k} g_i}{k\,t}$

2. Encode $\hat{\alpha}_a$

   Since t was previously encoded, only needs to encode $s_k = \sum_{i=1}^{k} g_i$

   Since $s_k \in \{0,1,...,t\}$, $\quad L(s_k) = \log(1+t)$

3. Encode $\mathbf{g} = \{g_1,...,g_N\}$ given $\hat{\alpha}_a$

   Should we plug $\hat{\alpha}_a$ in the multinomial ?

   Attention ! Incompleteness! Receiver already knows s_k

14 MDL for change-point detection in Poisson sequences

Best code for $\mathbf{g} = \{g_1,...,g_N\}$

…knowing that $\sum_{i=1}^{k} g_i = s_k$

..and consequently that $\sum_{i=k+1}^{N} g_i = t - s_k$

$$L(\mathbf{g} \mid t, m_k, s_k) = -\log \text{Multi}(g_1,...,g_k \mid s_k, \frac{1}{k},...,\frac{1}{k}) - \log \text{Multi}(g_{k+1},...,g_N \mid (t-s_k), \frac{1}{N-k},...,\frac{1}{N-k})$$

15 MDL for change-point detection in Poisson sequences

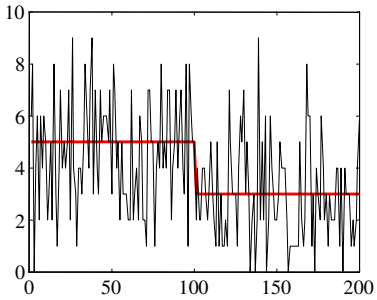
Finally, we choose the shortest of the N total description lengths

$$L_0 = L(\mathbf{g} \mid t, m_0)$$

$$L_k = \log(t+1) + L(\mathbf{g} \mid t, m_k, s_k), \quad k = 1,..,N-1$$

Interestingly: this MDL criterion coincides with the Bayesian model selection criterion with a uniform prior on α_k

16 MDL for change-point detection in Poisson sequences



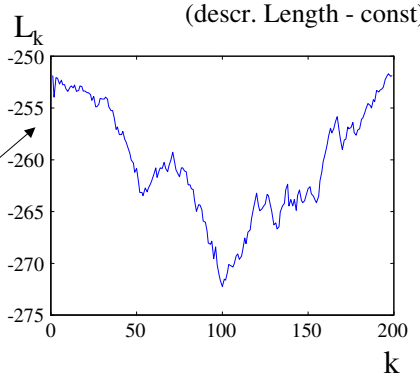
Change point at 100

$\lambda_a = 5$

$\lambda_b = 3$

Minimum at k = 100
the correct vale (in this case)

(descr. Length - const)

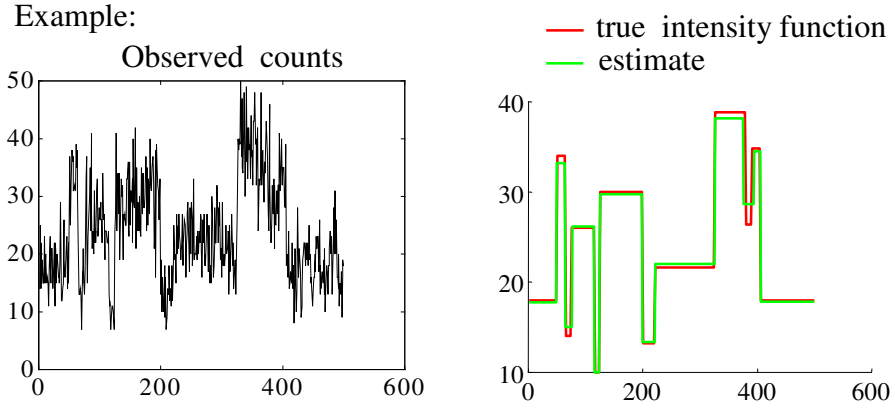


17 An MDL criterion to segment Poisson data

What about multiple change points ?

Simply take each segment and apply the criterion again (recursively).

Example:

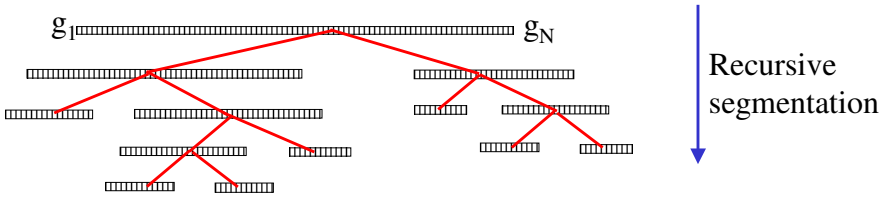


18 An MDL criterion to segment Poisson data

Recursive MDL segmentation

↓

Pyramidal (tree-structured) coding

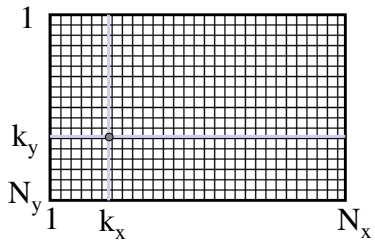


Notice: sending a given s_k corresponds to sending t, for the level below.

19 Segmenting Poisson images

Basic building block:

In 2D, look for the best (if any) segmentation into two/four rectangles:



Same multinomial-based criterion:

- condition on the total count
- number of competing models:

$$1 + N_x N_y$$

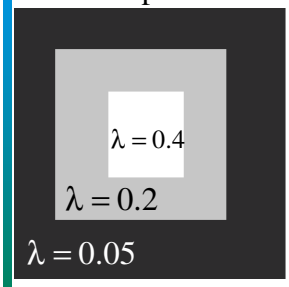
no segmentation,
all possible 4-segmentations,
and all possible 2-segmentations

To fully segment an image: apply criterion recursively.

20 Segmenting Poisson images: a synthetic example

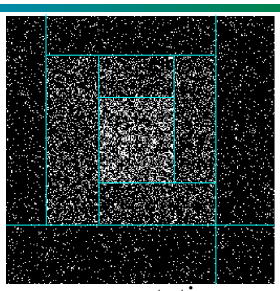
Example:

250*250 pixels/bins



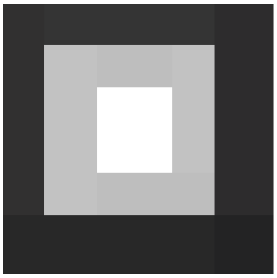
True intensity λ

counts



segmentation

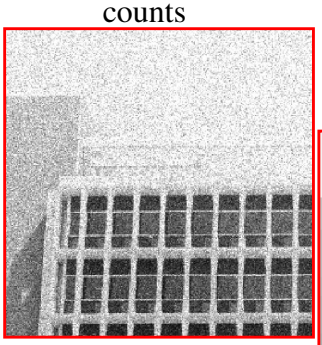
intensity estimate $\hat{\lambda}$



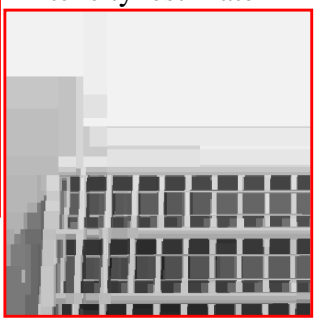
21 Segmenting Poisson images: another example

True intensity λ

counts



intensity estimate $\hat{\lambda}$



22 Conclusions

- Fully unsupervised segmentation of Poisson data
- …based on recursive application of MDL criterion.
- The MDL criterion (with insufficiency removed) is equivalent to a Bayes criterion.
- Corresponds to coarse-to-fine progressive coding

Future work

- Try an MDL-controlled merge phase
- Use more complex models for each segment (e.g., polynomial)
- Use non-rectangular partitioning in 2D