Array processing and statistical modelling for the detection and classification of seismic events

Paul Bui Quang
joint work with
Pierre Gaillard, Yoann Cano

CEA / DAM Ile-de-France
Laboratory of Detection and Geophysics
Low level signal processing: **detection** of any coherent wavefront crossing an array → many detections (including coherent noise)

Hence the need to **classify** detections thanks to higher level statistical modelling
**PMCC**: array processing for seismic (and infrasound) event monitoring

Progressive increase of array aperture

Estimation of parameters in elementary time-frequency domains

- propagation (speed, azimuth)
- energy (consistency)
**HMMs** are a powerful machine learning tool

Applications:

- speech recognition (Rabiner, 1989)
- seismic signal processing (Ohrnberger, 2001)

Beyreuther et al., 2008

Word “dry” (2.7s)

Earthquake (37s)
Hidden Markov models
Definition

HMM: time series statistical model

- \((X_t)_t\) Markov chain, hidden state of the system
- \(X_t\) takes discrete values \(\{1, \ldots, N\}\)
- \((Y_t)_t\) observation sequence
- likelihood: mapping \(\theta \mapsto \ell_\theta((Y_t)_t)\), where \(\ell_\theta(\cdot)\) is the observation density
Feature extraction

The raw signal is transformed into a sequence of observations

The observation sequence is modelled by a HMM

observation $Y_t = \text{vector of features}$
Let $a$ and $b$ be two event classes

Training data: $(Y_t^a)_t, (Y_t^b)_t$

Training consists in maximum likelihood estimation

$$\hat{\theta}^a = \arg\max_{\theta} \{\ell_{\theta}((Y_t^a)_t)\}$$
$$\hat{\theta}^b = \arg\max_{\theta} \{\ell_{\theta}((Y_t^b)_t)\}$$

Test data (unknown class): $(Y_t^?)_t$

Test event classified as:

- $a$ if $\ell_{\hat{\theta}^a}((Y_t^?)_t) > \ell_{\hat{\theta}^b}((Y_t^?)_t)$
- $b$ otherwise
Classification of seismic events
Study case

IMS seismic array station of Songino (Mongolia)

Classes:
• regional events
• teleseismic events
• noise

• 7821 PMCC detections
• 121 regional events
• 121 teleseismic events

*Test data*: September 1—2, 2008
• 3740 PMCC detections
• 23 regional events
• 53 teleseismic events

Event bulletin provided by RCAG
Classification of seismic events
Choice of features

Trajectory distribution of two discriminating features

spectral energy (log) 0.8—3Hz

PMCC consistency (log)
### Classification of seismic events

#### Experiment results

<table>
<thead>
<tr>
<th></th>
<th>Regional</th>
<th>Teleseismic</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional</td>
<td>93%</td>
<td>3%</td>
<td>4%</td>
</tr>
<tr>
<td>Teleseismic</td>
<td>2%</td>
<td>89%</td>
<td>8%</td>
</tr>
<tr>
<td>Noise</td>
<td>6%</td>
<td>4%</td>
<td>90%</td>
</tr>
</tbody>
</table>

**Training data**

<table>
<thead>
<tr>
<th></th>
<th>Regional</th>
<th>Teleseismic</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional</td>
<td>83%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>Teleseismic</td>
<td>2%</td>
<td>83%</td>
<td>15%</td>
</tr>
<tr>
<td>Noise</td>
<td>5%</td>
<td>9%</td>
<td>86%</td>
</tr>
</tbody>
</table>

**Test data**
Classification of seismic events
Experiment results (cont.)

Correct classification
False classification
Classification of seismic events
Experiment results (cont.)

Correct classification
False event classification
(regional <-> teleseismic)
False noise classification
Conclusion

classification scores: regional event, teleseismic event, noise

PMCC detection + HMM classification: promising association for automatic seismic analysis
Some references


• P. Bui Quang, P. Gaillard, Y. Cano, M. Ulzibat. Detection and classification of seismic events with progressive multi-channel correlation and hidden Markov models. [Submitted]