Processing Cardiotocographic Signals

The use of compression in cardiotocographic signal to detect fetus at risk.

Cristina Santos
csantos@med.up.pt
Cardiotocography

- **Cardiotocography**: recording the fetal heart rate (FHR) and uterine contractions (Toco).

- Widely used for fetal monitoring during labor.
Cardiotocography evaluate the fetal health condition and allow the obstetrician to intervene to prevent potential compromise and irreversible damage however with high inter-observer interpretation disagreement\textsuperscript{1-2}

The certainty-agreement diagram

The certainty-agreement diagram

Automated analysis: Omniview SisPorto

- automated approaches:

  Automatic estimation of:
  - uterine contractions
  - fetal heart rate baseline
  - identification of accelerations
  - identification of decelerations
  - quantification of short- and long-term variability
Automated analysis: Omniview SisPorto
Automated analysis

• Challenge:
  – signal processing and pattern recognized techniques to identify fetus at risk
Complexity

- **different approaches** based on nonlinear dynamics, chaos and complexity has been considered, which recognizes **irregularity, subjectivity** and **uncertainty** as intrinsic and fundamental

- **complexity** - quantifies the amount of structured information.
Entropy

• Shannon entropy: average **unpredictability** in a random variable

• With the goal of quantification of the **amount of regularity** in (heart rate) time-series data:
  – Pincus presented the Aproximate Entropy (ApEn) and later
  – Richman and J.R. Moorman presented Sample Entropy (SampEn)
Entropy

• A low ApEn value is associated with high degree of regularity.
  – a time series containing many repetitive patterns has a small ApEn;
  – a less predictable process has a higher ApEn.

• In the particular case of fetal heart rate (FHR) tracings, entropy measures have been widely used to detect different pathologies.
Compression

• However, we believe, that compression can be effectively applied as an alternative measure of complexity to the widely used entropy in biological signals.

• Compression is a measure of system complexity, but has been used to a lower extent in the analysis of biomedical signals.
Compression

• Kolmogorov complexity: the smallest representation of the object.

• A new clustering method **CompLearn** based on Kolmogorov Complexity, a well-studied notion of information content in individual objects, was introduced.


• You can try it in: [http://www.complelearn.org/](http://www.complelearn.org/)
Compression

• Good results were obtained using this approach in different areas, including:
  – literature\(^1\)
  – music\(^2\)
  – computer virus and internet traffic analysis\(^3\)

CompLearn

**Input:** a set of files

1. File translations, if necessary.
2. Calculation of Normalized Compression Distance (NCD).
3. Representation, unrooted binary tree.

**Output:** hierarchical clustering.
CompLearn

- This clustering system is unique in that it can be described as \textit{feature-free}.

- \textit{There are no parameters to tune}, and no domain-specific knowledge went into it.

- Using general-purpose data compressors gives us a parameterized family of features automatically for each domain.
CompLearn

\[ NCD(x, y) = \frac{C(x, y) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}} \]

Computing NCD of every FHR with every other FHR yields a 2-dimensional symmetric \textbf{distance matrix}.

Next step is transforming this array of distances into something easier to grasp.

Use the \textbf{Quartet Method} to construct an unrooted binary \textbf{tree} from the NCD matrix.
In 2006 we applied to cardiotocographic tracings

Learning by Compression

Keogh showed that when clustering heterogeneous data and anomaly detection in *time sequences*, the compression approach outperforms every known data-mining method.

PAQ8L compressor

• PAQ8 represents a series of lossless compressors with the world’s highest compression ratio. PAQ8L, based on Dynamic Markov compression, was released in 2007.

• We believe that these compressors can be successfully used in the medical field as well.
Entropy versus compression

• 68 cardiotocographic tracings:
  – 48 delivered fetuses with umbilical artery pH in the normal (N) range
  – 10 delivered fetuses with umbilical arterial pH mildly acidemic (MA) fetuses
  – 10 delivered fetuses with umbilical arterial pH moderate-to-severe acidemic (MSA) fetuses

All traces were resampled at a frequency of 2Hz
Entropy versus compression

Median, first ($Q_1$) and third($Q_3$) quartiles of complexity measures of fetal heart rate (FHR) tracings from moderate-to-severe acidemic (MSA), mildly acidemic (MA) and normal (N) fetuses in the **final 5 min segments**.

<table>
<thead>
<tr>
<th></th>
<th><strong>MSA</strong></th>
<th></th>
<th><strong>MA</strong></th>
<th></th>
<th><strong>N</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>($Q_1$, $Q_3$)</td>
<td>Median</td>
<td>($Q_1$, $Q_3$)</td>
<td>Median</td>
<td>($Q_1$, $Q_3$)</td>
</tr>
<tr>
<td>Entropy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ApEn(2,0.1)</td>
<td>0.585</td>
<td>(0.525, 0.733)</td>
<td>0.738</td>
<td>(0.686, 0.774)</td>
<td>0.682</td>
<td>(0.555, 0.739)</td>
</tr>
<tr>
<td>ApEn(2,0.15)</td>
<td>0.496</td>
<td>(0.291, 0.738)</td>
<td>0.642</td>
<td>(0.561, 0.792)</td>
<td>0.607</td>
<td>(0.490, 0.702)</td>
</tr>
<tr>
<td>ApEn(2,0.2)</td>
<td>0.351</td>
<td>(0.251, 0.553)</td>
<td>0.582</td>
<td>(0.469, 0.795)</td>
<td>0.516</td>
<td>(0.420, 0.627)</td>
</tr>
<tr>
<td>SampEn(2,0.1)</td>
<td>0.476</td>
<td>(0.325, 0.658)</td>
<td>0.598</td>
<td>(0.540, 0.985)</td>
<td>0.541</td>
<td>(0.402, 0.615)</td>
</tr>
<tr>
<td>SampEn(2,0.15)</td>
<td>0.309</td>
<td>(0.172, 0.636)</td>
<td>0.459</td>
<td>(0.403, 0.632)</td>
<td>0.434</td>
<td>(0.320, 0.549)</td>
</tr>
<tr>
<td>SampEn(2,0.2)</td>
<td>0.231</td>
<td>(0.172, 0.307)</td>
<td>0.369</td>
<td>(0.308, 0.637)</td>
<td>0.341</td>
<td>(0.256, 0.404)</td>
</tr>
<tr>
<td>Compression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>paq8l</td>
<td>234.0</td>
<td>(211.0, 279.0)</td>
<td>355.0</td>
<td>(306.0, 393.0)</td>
<td>335.0</td>
<td>(293.5, 372.5)</td>
</tr>
<tr>
<td>bzip2</td>
<td>283.5</td>
<td>(270.0, 382.0)</td>
<td>444.0</td>
<td>(404.0, 501.0)</td>
<td>426.5</td>
<td>(362.5, 488.0)</td>
</tr>
</tbody>
</table>
Entropy versus compression

Scatter plot of indices ApEn(2,0.15) and paq8l for the final 5 min segments, comparing normal fetuses (*), mildly acidemic fetuses (circle) and moderate–severe academic (MSA) fetuses (◊).
Conclusions

• Compression - rarely been used
• Compressors can be an alternative to the widely used entropy metrics
• Allows the computation in real time – smaller running times
• Low correlation – combination of the entropy and compression