ENF SIGNAL EXTRACTION FOR ROLLING-SHUTTER VIDEOS USING PERIODIC ZERO-PADDING

Jisoo Choi and Chau-Wai Wong
North Carolina State University, Raleigh, USA

Background & Motivation

- **Electric Network Frequency (ENF)** is the supply frequency of electric power grid. 60 Hz in North America, 50 Hz in most other regions.
- ENF fluctuates around nominal due to mismatch between demand and supply.
- Define instantaneous values of ENF over time as the **ENF Signal**.
  - How to extract? Short-Time Fourier Transform; Subspace methods, e.g., MUSIC, ESPRIT.
  - **Natural timestamp**: Inherently embedded in audio and visual recordings via sensing acoustic vibrations/electromagnetic interference and near-invisible flickering of lighting.
  - Promising forensic techniques: multimedia authentication and tampering detection.

Proposed Method for ENF Extraction From Videos

- **Rolling shutter** acquires video by sensing rows sequentially. Periodic gaps exist between frames.
- “Direct concatenation method” [1]: Concatenates samples obtained by rolling shutter and ignores periodic gaps.
- Proposed **periodic zeroing-out method**: Treats as missing value problem and uses zeros for idle period at end of each frame. Multirate analysis was conducted.
- Theoretical result: (1) The periodically zeroed out version of the ideal signal \( x(n) \) is the weighted sum of shifted \( X(F; F_k) \)'s:
  \[
  Y(F; F_k) = \sum_{m=-\infty}^{M-1} A_m X(F - mF_s; F_k)
  \]
  where \( F_k = \frac{F_s}{T} \) and \( A_m = \frac{1}{M} \sin(\frac{2\pi m}{M}) \cdot \exp(-j\frac{2\pi m}{M}+1) \).
- (2) Strongest ENF traces are always found at 2× 60/50 Hz.

Experimental Results

- Aliased ENF signals and aliased DC signals appear consistent with theoretical prediction.
- Extraction ENF signal is of high accuracy when compared to reference signal.

Conclusions & Future Work

- Periodic zeroing-out method for ENF extraction from rolling shutter acquired videos is proposed.
  - Extracted ENF signals are less distorted.
  - ENF trace with highest SNR is always located at 2× 60/50 Hz.
- To extensively evaluate on large datasets with different camera models and SNR conditions.

![Graph showing ENF extraction results](image)

**Pros and cons of the proposed method and the prior art.**

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>Average of correlations</th>
<th>Direct concatenation</th>
<th>Proposed</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
<td>0.959</td>
<td>0.951</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>-20</td>
<td>0.858</td>
<td>0.892</td>
<td>0.01**</td>
<td></td>
</tr>
<tr>
<td>-23</td>
<td>0.754</td>
<td>0.839</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>-25</td>
<td>0.459</td>
<td>0.551</td>
<td>0.01**</td>
<td></td>
</tr>
</tbody>
</table>