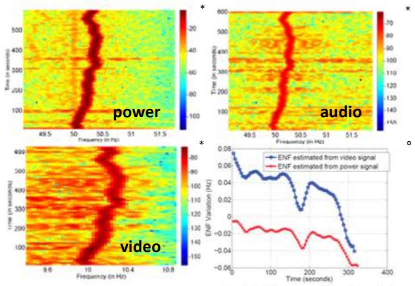
ENF SIGNAL EXTRACTION FOR ROLLING-SHUTTER VIDEOS USING PERIODIC ZERO-PADDING

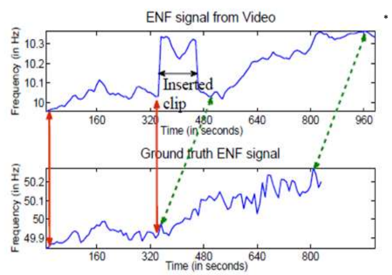
Jisoo Choi and Chau-Wai Wong
North Carolina State University, Raleigh, USA

NC STATE UNIVERSITY

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.





* R. Garg, A. L. Varna, and M. Wu, "'Seeing' ENF: Natural time stamp for digital video via optical sensing and signal processing," Proceedings of the 19th ACM International Conference on Multimedia, 2011.
o H. Su, A. Hajj-Ahmad, C.-W. Wong, R. Garg, and M. Wu, "ENF signal induced by power grid: A new modality for video synchronization," in ACM International Workshop on Immersive Media Experiences, 2014.

Proposed Method for ENF Extraction From Videos

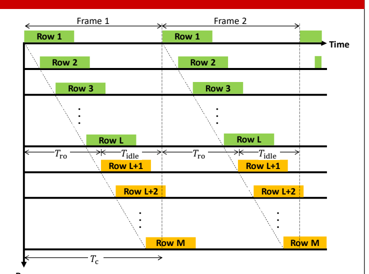
- ***Rolling shutter*** acquires video by sensing rows sequentially. Periodic gaps exist between frames.
- "Direct concatenation method" [1]: Concatenates samples obtained by rolling shutter and ignores periodic gaps.
- Proposed "***periodic zeroing-out method***": Treats as missing value problem and 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 $X(F; F_s)'s$:

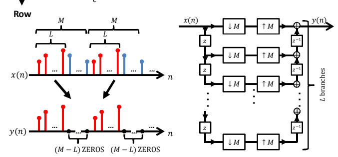
$$Y(F;F_s) = \sum_{m=0}^{M-1} A_m \, X(F - mF_c; F_s),$$

(Frame rate; Sampling rate)

$$\text{where } g(F;F_s) \stackrel{\text{def}}{=} g\left(\frac{2\pi F}{F_s}\right), A_m = \frac{L}{M}\text{asinc}_L\left(\frac{2\pi m}{M}\right) \cdot \exp\left(-j\frac{\pi m}{M}(L-1)\right).$$

(2) Strongest ENF traces are always found at 2× 60/50 Hz.

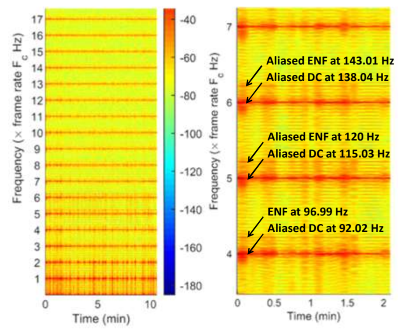




[1] H. Su, A. Hajj-Ahmad, R. Garg, and M. Wu, "Exploiting rolling shutter for ENF signal extraction from video," in IEEE International Conference on Image Processing, Oct. 2014, pp. 5367–5371.

Experimental Results

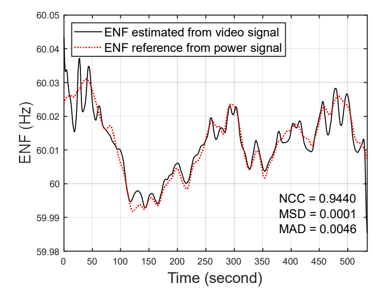
Aliased ENF signals and aliased DC signals appear consistent with theoretical prediction.



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

	Direct concatenation	Periodic zeroing-out (proposed)
Scaling function	Distorts power spectrum	Uniformly scales power spectrum
Freq. of highest SNR	Depends on T_c and T_{ro}	Always same at 2× nominal ENF
Need of knowing T_{ro}	Not needed if one does not care about locations of highest SNR component	Needed for generating row signal
Averaged noise	Larger than periodic zeroing-out method	Smaller than direct concatenation method

Extracted ENF signal is of high accuracy when compared to reference signal.



Under low SNR conditions, proposed extraction method can generate estimated ENF signals closer to ground-truth ENF signal than direct concatenation method does.

SNR (dB)	Average of correlations: Direct concatenation	Average of correlations: Proposed	*p*-value
-10	0.959	0.951	0.20
-20	0.858	0.892	0.01**
-23	0.754	0.839	0.18
-25	0.459	0.551	0.01**

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.