Automatic Search and Classification of Sound Sources in Long-Term Surveillance Recordings

Robert C. Maher and Joseph Studniarz

Electrical and Computer Engineering

Montana State University - Bozeman

Outline

- Introduction
- Experimental Procedure
  - Example recordings
  - Feature detection
- Experimental Results
  - Time-variant spectra
  - Identifying events
  - Effects of temporal and spectral overlap
- Discussion and Conclusions

Introduction

- Long-term audio recordings
  - Continuous recordings of urban areas and natural soundscapes
  - Duration may be days, weeks, months, …
  - Human audition infeasible due to length
- Forensic applications
  - Detection and Classification
  - Event reconstruction
  - Timeline Assessment

Introduction (cont.)

- Source may be deliberate surveillance or inadvertent surveillance
  - Wildlife studies
  - Regulatory monitoring
  - Acoustic surveys
  - Electronic newsgathering
  - Amateur A/V recordings

Introduction (cont.)

- Prior research in speech/non-speech segmentation, music recognition, and computational auditory scene analysis
- General research questions:
  - Search for a target sound, or classify all sounds?
  - Will conventional pattern matching techniques do the job?
  - What level of performance is necessary?

Experimental Procedure

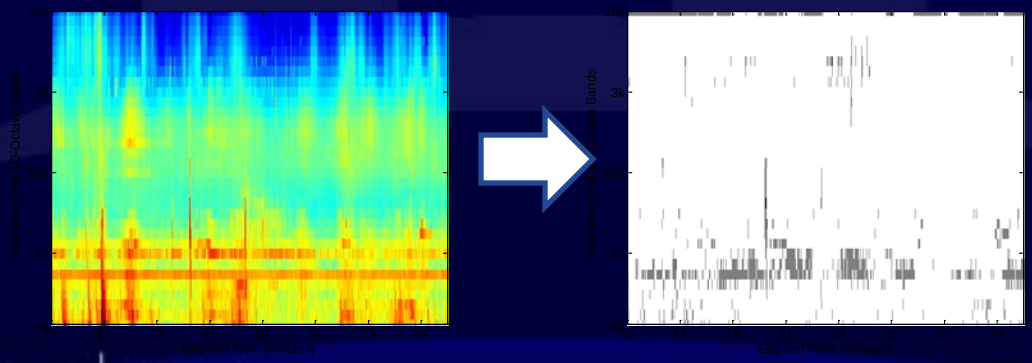
- Our example recordings
  - Grant-Kohrs Ranch (2009-10): 8,700 hours
  - Nyack River experiment (2012): 29 hours
- Sound composition
  - Biophony—birds, frogs, insects, mammals
  - Geophony—wind, rain, hail, thunder, waterfall
  - Anthrophony—aircraft, automobiles, trains, domestic sounds

Experimental Proc. (cont.)

- Typical recordings contain a background texture punctuated by foreground sounds—but also gradual evolution and overlaps in time and frequency.
- Isolated pattern matching not applicable in general.
- Current compromise: use automated search to identify sections of interest for subsequent audition

Technical Approach

- Treat spectrographic information as an *image*, or as a 3-D *surface*
- Sound events comprise *edges*, *ridges*, *cliffs*, and *valleys* in the spectrogram
- Plan: process the spectrographic information to reveal the feature space



Spectral-domain Filters

- Vertical edges = broad in frequency, narrow in time: clicks
- Horizontal edges = narrow in frequency, broad in time: tones
- Diagonal edges = tonal sweeps
- “Waffle” pattern viewpoint

Forensic Example

- One hour audio recording
  - Question 1: Is the sound of a gunshot present at any time in the recording?

  If the answer to question 1 is *yes,*

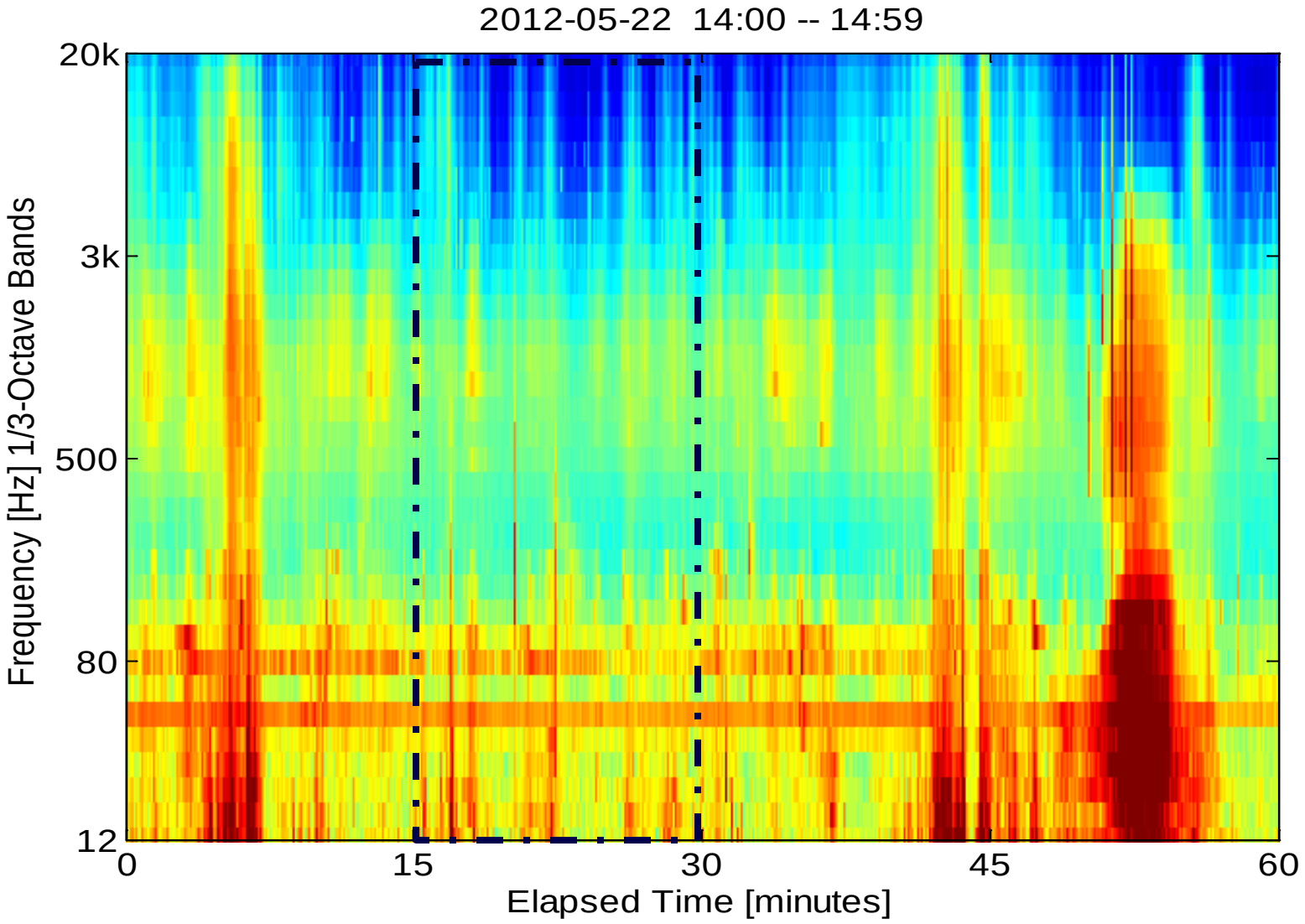
  - Question 2: Is there more than one recorded shot?

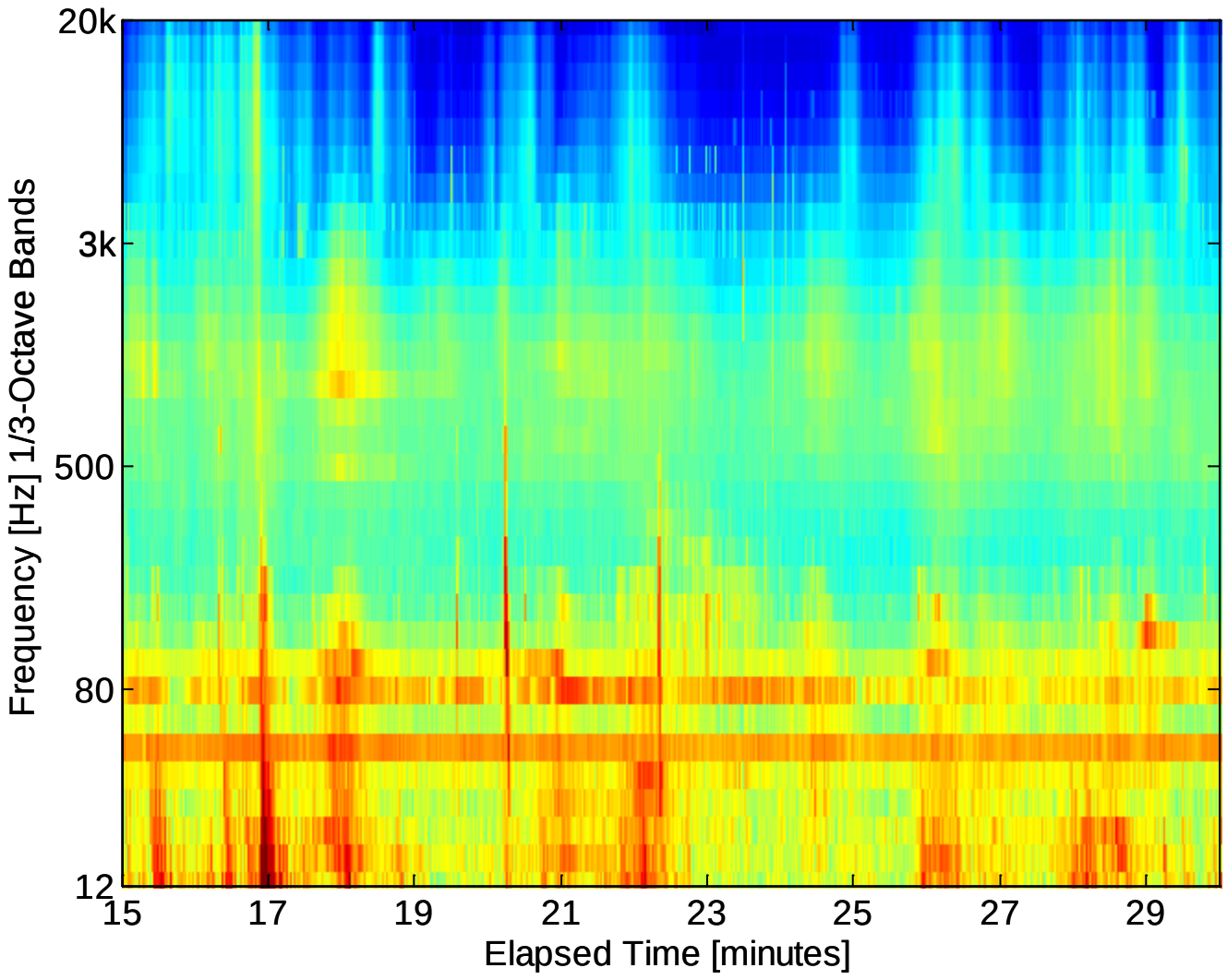
  If the answer to question 2 is *yes*,

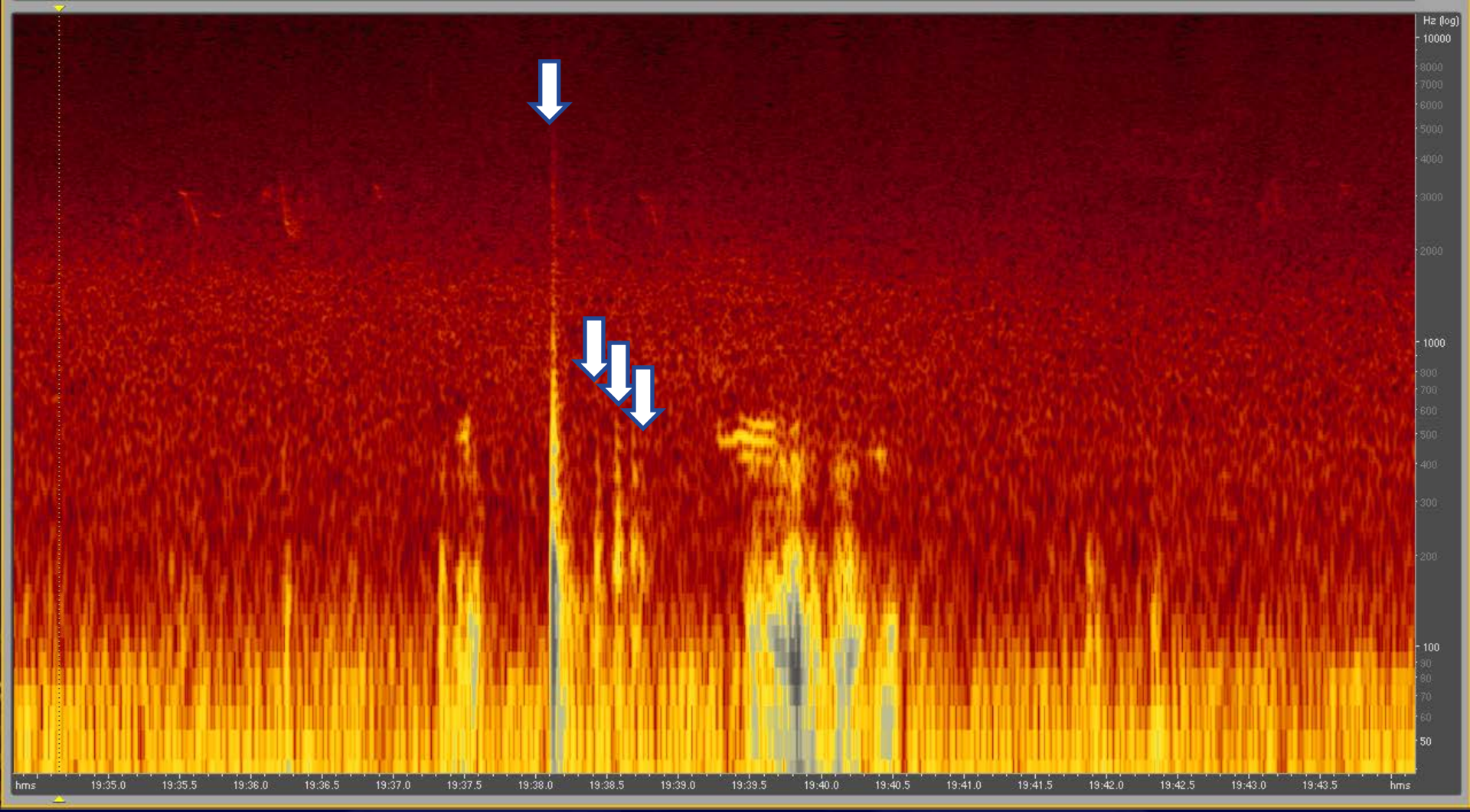
  - Question 3: Are the multiple shots from the same or from different firearms?

Forensic Example (an aside)

- This is a “needle in the haystack” analogy
- Finding the needle may be extremely difficult
- If no needle is found, it may mean:
  - no needle was present, or
  - the search process missed it
- But if a needle IS found, it will generally not be misinterpreted as some other object


2012-05-22 14:00 -- 14:59
20k
3k
500
80
12
Frequency [Hz] 1/3-Octave Bands
0
15
30
45
60
Elapsed Time [minutes]


20k
3k
500
80
12
Frequency [Hz] 1/3-Octave Bands
15
17
19
21
23
25
27
29
Elapsed Time [minutes]


Hz (log)
10000
8000
7000
6000
5000
4000
3000
2000
1000
800
700
600
500
400
300
200
100
90
80
70
60
50
hms
19:35.0
19:35.5
19:36.0
19:36.5
19:37.0
19:37.5
19:38.0
19:38.5
19:39.0
19:39.5
19:40.0
19:40.5
19:41.0
19:41.5
19:42.0
19:42.5
19:43.0
19:43.5
hms

Experimental Results (cont.)

- Edge-based spectro-temporal search can be implemented easily
- Good for abrupt sound events or narrow-band tones
- Sounds with spectral and temporal overlap are not handled well

Future Work

- Extend approach to detailed patterns, such as birdcalls and other elements of the biophony
- Establish reliability of using MP3 spectral frame data
- Classification algorithm development

Acknowledgements

- **Co-author Joe Studniarz (undergraduate research assistant at Montana State Univ.)**
- **NPS Natural Sounds and Night Skies Program, Ft. Collins, CO**
  - Kurt Fristrup, Emma Lynch, Damon Joyce
- Montana Institute on Ecosystems
  - Cathy Whitlock and Ric Hauer, Directors