Automated analysis and interpretation of long-term soundscape audio recordings

Robert C. Maher
Electrical and Computer Engineering
Montana State University - Bozeman
Outline

• Introduction
• Long-term acoustical acquisition
  – Soundscape studies (thousands and thousands of hours)
  – Environmental monitoring
• How to present soundscape information?
  – Spectrograms
  – Aural snapshots
  – Time-lapse aural presentation
• Conclusion
Introduction

• Long-term soundscape studies are now feasible and desirable
• Interpretation and presentation is difficult due to extreme length of the data
• What is needed:
  – Automated analysis tools
  – Useful comparison metrics
  – Meaningful presentation techniques
Example: Grant-Kohrs Ranch National Historic Site

- Deer Lodge, Montana
- A working cattle ranch commemorating the heritage of American cowboys, stock growers, and cattle operations during the 19th and 20th centuries.
- Congress: established in 1977 to maintain the site as a working ranch.
- Cultural soundscape is essential: all the sights, sounds, and sensations associated with ranching.
Long-Term Collection

March 17, 2009

June 22, 2009

September 5, 2009

December 12, 2009
Project Presentation

Challenges

• Audio recording lasting 365 days = 8,760 hours (525,600 minutes)
• Long segments of natural quiet with sections of recognizable biophony, geophony, and anthrophony
• Visitors to a web site or visitor center spend only a few minutes: can we compress meaningfully by 1/200,000?
Some options

• Automated SPL min/max/average graphs
• Spectrographic displays
• Audio samples of “highlights”
• Time-lapse aural display
SPL Graphs

• Presents information on maximum, minimum, and average sound levels
• Relatively simple to produce
• Interpretation still required
• Little information in general about sound sources and distributions
SPL Graphs

Jul-09

Sound Pressure Level [dBA re 20 uPa]

Elapsed Hours
Spectrographic Display

- Conveys time-frequency-energy distribution
- Condenses a lot of information into a compact form
- May be confusing to the public unless explained
- Works best if audio playback allowed (point and click)
Audio highlights excerpts

• Identify and extract “interesting” sound examples
• The visitor can quickly sample the range of sounds and sound textures
• Generally requires considerable audition and manual preparation
• May give a non-representative indication of the actual sound texture
Time-lapse aural display

- **Goal:** represent the aural sound texture for many minutes of real time audio with only a few seconds of seamless excerpts
  - Aural equivalent of time-lapse photography
- **Challenge:** defining and capturing sound texture in an aurally meaningful manner
  - Simple block-downsampling may not capture sonic *texture* effectively
Non-uniform time warp concept

Compressed Time

\[ \frac{T_s}{N} \]

Normal Time

\( t_0 \quad t_1 \quad t_2 \quad t_3 \quad t_4 \quad t_5 \quad T_s \)
Approach for time-lapse aural display

• Create a spectral transition map: identify textural boundaries in the audio
• Determine available segments based on compression factor $N$
• Assign segments to the transitions in order of priority
• Segment the audio and concatenate with overlap-add
Example Transition Map (N=10)

8kHz

0

2' 10"
Reconstructed Signal (N=10)
Conclusion

• Long-term acoustical acquisition requires automated analysis and distillation tools
• Presenting days/weeks/months of audio is challenging
• Extreme time-scale compression is necessary for many applications
• Ongoing effort is needed in pattern detection and pattern matching
Acknowledgements

- **Grant-Kohrs Ranch NHS, Deer Lodge, MT**
  - Christine Ford, Integrated Resources Program Manager

- **NPS Natural Sounds Program, Ft. Collins, CO**
  - Kurt Fristrup, Emma Lynch, Damon Joyce

- **Rocky Mountain Cooperative Ecosystem Studies Unit (RM-CESU), Missoula, MT**
  - Kathy Tonnessen, Natural Resources Research Coordinator
  - Lisa Gerloff, Executive Coordinator
Sound Examples

- March 18, 2009 9:34PM MDT (45”)
- April 15, 2009 6:13AM MDT (before dawn) (1’)
- May 1, 2009 11:22AM MDT (5’)
- May 4, 2009 6:23AM MDT (after dawn) (2.5’)
- July 6, 2009 ~noon (6.5’)
- Dec 30, 2009 9:30PM MST (2.5’)
- July 6, 2009 ~1:30PM MDT (2’)

- http://ece.montana.edu/rmaher/audio_monitor/grko.htm