GENERAL-PURPOSE REAL-TIME MONITORING OF MACHINE SOUNDS

Stephen V. Rice
The University of Mississippi
Dept. of Computer and Information Science
P.O. Box 1848
University, MS 38677 USA
rice@cs.olemiss.edu

Stephen M. Bailey
Comparisonics Corporation
P.O. Box 1960
Grass Valley, CA 95945 USA
sbailey@comparisonics.com

Abstract: The "Comparisonics" sound-matching algorithm computes signatures to characterize audio data and compares the signatures to measure the similarity of sounds. This algorithm is general purpose: it can compare any sounds and can monitor sounds from any machine. After a baseline recording is characterized, signatures are derived from the current sound and compared in real time with the baseline signatures. Similarity scores are computed continuously, reflecting the similarity of the current sound to the baseline, and an alert can be given if the similarity falls below a threshold. In addition, the current sound can be compared with known "error" sounds to identify specific faults. This algorithm can operate in a wireless sensor network and be deployed in a factory or on a ship to monitor an unlimited variety of machinery.

Key Words: Acoustic emissions; audio comparison; condition monitoring; wireless sensor networks.

Introduction: A human listener accustomed to the "normal" sound of a machine readily detects changes in the sound. It is well known that the sounds made by machines can indicate the health of the machinery. A change in sound may portend trouble and warrant investigation by skilled maintenance personnel. However, a change may go unnoticed by human listeners due to unfamiliarity with the usual sound, or because the change is gradual, sporadic, or masked by the din of neighboring machinery. If the machine is remote and unsupervised, there is no one to listen.

There is a great need for computers to listen to machinery and many millions of machines could be usefully monitored; however, a robust low-cost solution has been unavailable. Let us consider some existing approaches:

- *amplitude monitoring* – a simple device detects whether the sound is louder or softer than expected; it is a low-cost, general-purpose solution but it does not perform any frequency analysis; it detects changes in the quantity, but not the quality, of sound

- *machine-dependent pattern recognition* – a custom pattern-recognition system is developed to monitor the sounds from a particular machine; accurate monitoring may be achieved for this machine but the cost to develop the system is high

- *machine-class-dependent pattern recognition* – characteristics of a particular class of machines (for example, rotating machinery) are exploited to monitor machines belonging to the class; the cost is high but not as high as developing a custom pattern-recognition system for each machine in the class

The current situation, as we see it, is a choice between low-cost amplitude monitoring and high-cost pattern recognition. The semi-automated solution in which trained humans analyze FFT output belongs to the latter category.

We have developed a solution that might be considered "machine-independent pattern recognition." Its cornerstone is a general-purpose "sound-matching" algorithm that compares any sounds and measures their similarity. This algorithm can listen to and compare sounds from any type of machine. After characterizing a baseline recording of a machine, it compares the current sound with the baseline to detect changes in real time.

The algorithm is ideally deployed in a wireless sensor network, where each wireless node is equipped with a microphone. The wireless nodes can be easily placed by personnel throughout a factory or ship, near machinery to be monitored. Although the algorithm can compare vibration data, we avoid the use of accelerometers, which are more expensive and more difficult to install than microphones. After some initial configuration, the system begins real-time monitoring of sounds.

By employing a general-purpose sound-comparison algorithm and inexpensive, easy-to-install wireless nodes with microphones, the goal of robust, low-cost sound monitoring can be achieved.

Sound Matching: The Comparisonics sound-matching algorithm was developed by S. V. Rice in 1997. The initial goal was to develop a method for "content-based" retrieval of sound effects. Previously, sound-effect collections could be searched only by entering a text description for each sound and then performing a keyword search of the text descriptions. In Rice's "sounds-like search," a user can specify any example sound and the system automatically retrieves perceptually similar sounds. This unique search capability has been incorporated into FindSounds.com, the first Web search engine for sound effects [1,2].

The sound-matching algorithm processes the uncompressed sequence of sample values in any digital audio recording, provided the duration of the recording is at least ten milliseconds and the sample rate is at least 8 kHz. The algorithm characterizes the sounds in the recording by a "signature," which is a 16-byte quantity that encodes a vector of perceptual features. Given any two signatures, the algorithm returns a score ranging from 0 (least similar) to 100 (most similar) describing the similarity of the recordings from which the signatures are derived. In assigning a score, the algorithm emulates the human perception of sound similarity: the higher the score, the more a human listener will perceive the two recordings to be alike. Nearly all of the perceptual features are characterizations of frequency content and are extracted from the time domain by a proprietary transform. A nonlinear distance measure in the multi-dimensional feature space is used to compute similarity scores.

The sound-matching algorithm was designed to produce a meaningful measure of sound similarity for all audible sounds. It is general purpose, not trained to a particular class of sounds. For sound-effects retrieval, it is impractical to develop a custom algorithm for each type of sound effect: rain, sirens, elephants, etc. Likewise, we cannot afford to build a custom algorithm for each machine.

The efficiency of the Comparisonics algorithm is notable. On a modern personal computer, the time required to compute a signature is less than one percent of the duration of the recording; thus, signatures can easily be computed in real time. The time required to compare two signatures and compute a similarity score is less than 0.5 microseconds. A signature occupies only 16 bytes for efficient manipulation and storage. (Contrast this with the storage required for FFT output, which is measured in kilobytes.)

One signature can be computed for a long recording and represents the "average" of the sounds in the recording. For machinery monitoring, it is desirable to divide the recording into consecutive 100-millisecond intervals and compute one signature to characterize each interval. This sequence of signatures can be compared with another sequence of signatures if the temporal ordering of sounds is to be matched. However, in our standard implementation, we ignore the temporal sequence.

Let $S = \{s_1, s_2, \ldots, s_n\}$ be an unordered set of n signatures derived from one or more baseline recordings of a machine. We apply a clustering algorithm to find a set of signatures $C = \{c_1, c_2, \ldots, c_m\} \subseteq S$ representing the distinct sounds in the recording, where each c_i is the exemplar of a cluster of perceptually similar sounds. In this way, the set C characterizes the natural variations of the "normal" sound of the machine. These variations may occur in one operating mode or be produced by different modes (for example, the cycles of a washing machine).

To monitor the sound of the machine in real time, a signature s is computed for the current 100-millisecond interval and the similarity of s and c_i is computed for each $i = 1, 2, \ldots, m$. The maximum similarity score represents the best match between the current sound and the baseline recording. If the maximum score falls below the value of

a threshold parameter τ, then an alert can be raised immediately, or the alert might be issued only after the maximum score has fallen below the threshold for multiple time intervals. In Fig. 1, the maximum score is plotted versus time for a recording of a compressor. The score declines as the sound of the compressor deviates from the baseline recording.
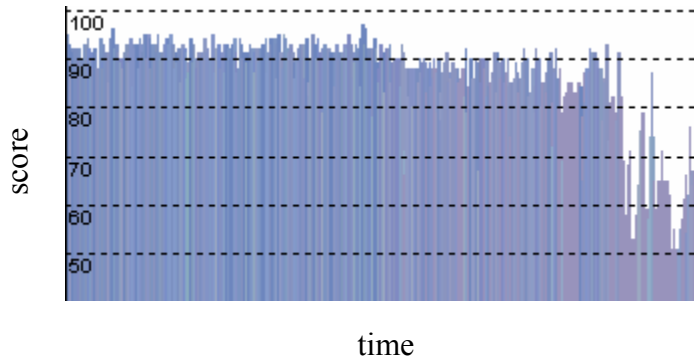

Fig. 1. Monitoring the Sound of a Compressor

The choice of τ can be determined manually through experimentation or automatically by comparing the signatures derived from one baseline recording with the exemplars of another baseline recording. In the latter approach, τ is chosen to be less than the observed scores.

The algorithm currently does not factor in the relative "strengths" of the clusters. That is, some clusters represent more sounds in the baseline recording than others. However, this can be considered to determine whether the current sound is rare or common in the baseline recording.

The algorithm can characterize not only baseline recordings but also recordings of known "error" sounds. If the current sound deviates from the baseline, it can be compared with the signatures of faults in an attempt to diagnose the problem.

Testing: The robustness of the Comparisonics sound-matching algorithm has been demonstrated for a wide variety of sounds in its application to sound-effects retrieval. FindSounds.com is utilized by more than 150,000 users per month.

A collection of recordings of 600 different machines has been assembled from sound-effects compilations for testing the machinery monitoring application. The sound-matching algorithm was not modified or specially trained for this test, yet it reliably matches the sounds in this collection. The diversity of this collection demonstrates the general-purpose nature of the algorithm. Below are some of the sound sources in this collection:

- acetylene torch
- air conditioner
- anesthesia ventilator
- arc welder
- band saw
- boiler
- cement mixer
- centrifuge
- clothes dryer
- clothes washer
- compressor
- conveyor belt
- drill press
- fan
- furnace
- gears
- generator
- grinder
- lathe
- microwave oven
- milling machine
- motor
- oxygen mask
- packing machine
- photocopier
- pile driver
- planer
- printing press
- pulse oximeter
- pump
- respirator
- stamping machine
- steel cutter
- table saw
- threshing machine
- transformer
- turbine engine
- winch

In addition, the test collection includes sounds from vehicles: airplane, bus, car, elevator, ferry, helicopter, motorcycle, rollercoaster, ship, subway, tank, tractor, train, and truck.

Proof of concept has been demonstrated using personal computers. However, it is impractical to place a PC next to each machine to monitor. The next step is to port the sound-matching algorithm to a wireless sensor platform and conduct tests of the algorithm on this platform. A project is underway to implement and test the algorithm on Crossbow Technology's MICA Mote platform [3].

Conclusion: Recent issues of the *Communications of the ACM* and *IEEE Computer* are devoted to the burgeoning field of wireless sensor networks [4,5]. Technologies for monitoring are rapidly evolving. For machinery monitoring, comparing audio and vibration signals is a challenge. A general-purpose sound-matching algorithm avoids the costs and complexities of custom development, and deployed on a wireless platform, provides a cost-effective solution for monitoring the sounds of an unlimited variety of machinery.

References:
1. S. V. Rice and S. M. Bailey, "Searching for Sounds: A Demonstration of FindSounds.com and FindSounds Palette," *Proceedings of the International Computer Music Conference*, Coral Gables, Florida, November 2004, pp. 215-218.
2. http://www.FindSounds.com
3. http://www.xbow.com
4. *Communications of the ACM*, vol. 47, no. 6, June 2004.
5. *IEEE Computer*, vol. 37, no. 8, August 2004.