inside the cooler was studied in order to determine its effects on the cooling power and to minimize its effects by modifying the cooler geometry. The streaming effect deals with acoustically stimulated gas flow within the cooler which transfers heat to its cold parts; this is due to acoustically stimulated viscous losses with the cooler walls and with the stack which causes forced connection .

As a result of this study dealing with higher acoustic intensity levels we show that an optimized high frequency thermoacoustic cooler can reach cooling power density levels of a few hundred watts per square centimeter, with applications to laser cooling, electronic cooling, and other applications needing thermal management.

This work is funded by The Office of Naval Research.

This article is based on paper 4pPA11 presented at the 151st ASA meeting in Providence.

Husam El-Gendy, Laurence Lyard, and Orest Symko.

CLASSIFYING KILLER WHALE VOCALIZATION USING TIME WARPING

Judith C. Brown
Physics Department, Wellesley College
Wellesley, Massachusetts 02481 and
Media Laboratory, Massachusetts Institute of Technology
Cambridge Massachusetts 02139

Patrick J. O. Miller
Sea Mammal Research Unit, University of St. Andrews
St. Andrews, Fife KY16 9QQ, UK

Fig. 1 Recording killer whale sounds.

Introduction and background

Marine mammals produce a wide range of vocalizations; therefore an improved ability to classify recorded sounds could aid in species identification as well as in tracking movements of animal groups. Killer whales produce three forms of vocalizations: clicks, whistles, and pulsed calls. Clicks consist of an impulse train (series of broadband pulses); whistles consist, for the most part, of a single sinusoid with varying frequency; and pulsed calls are more complex sounds with many harmonics. The repetition rate is a measure of the periodicity of the signal, and its measurement is called “pitch tracking” or “fundamental frequency tracking” in the speech literature. This measure gives rise to a pitch contour which is equivalent to the melodic line of a song (see Fig. 2).

Killer whales produce a number of highly stereo-

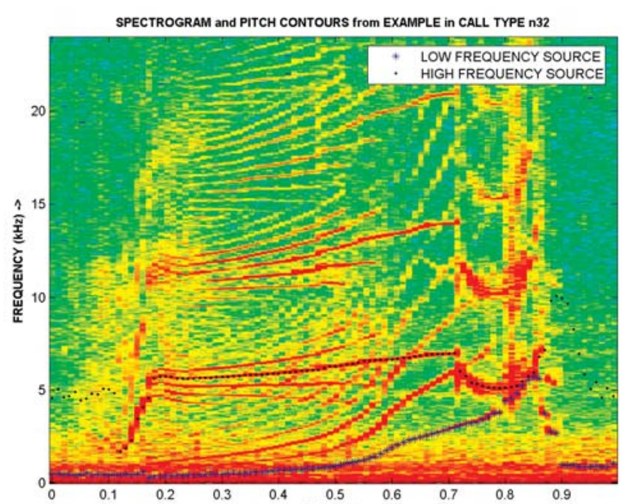

Fig. 2. Spectrogram showing pitch contours of the low frequency and high frequency sources in one killer whale pulsed call. Note there is noise before and after the onset of the calls.

Continued on page 46

Continued from page 45

typed (repeated and recognizable) pulsed calls, which are thought to be learned within the pod (living group). Repertoires of these stereotyped calls are pod specific, and the pitch contours of shared stereotyped calls are also group-specific from matrilineal lines (group with same mother) to larger pods (consisting of several matrilineal lines) to clans (even larger groups sharing calls). One of the remarkable features of killer whale pulsed calls is that they contain two overlapping but independently modulated contours or "voices." These are shown superposed on the spectrum as in Fig. 2. Bi-phonation, as this is called, is common in birds but has been described for few other marine mammal sounds. One of the challenges of analyzing these complex sounds is to "pitch-track" these two components from the same sound as shown in the example.

For the most part, the sounds produced by killer whales have been classified into groups called "call types" by humans from listening to the calls and observing their spectra.[1] This human classification by eye and ear is quite consistent, and has been useful to reveal group-specific acoustic repertoires and matching vocal exchanges. It would, nonetheless, be useful to replace human classification with an automatic technique because of the large amounts of data to be classified, and the fact that automatic methods can be fully replicated in subsequent studies.

In our studies we examined two sets of sounds previously classified into call types by human listeners. The first set was recorded from captive killer whales in Marineland in the French Antilles, and the second set from northern resident whales recorded on the open sea.

Dynamic time warping (DTW) and dissimilarity of pitch contours

The sounds that were classified into each call type have a similar shape or contour within that group

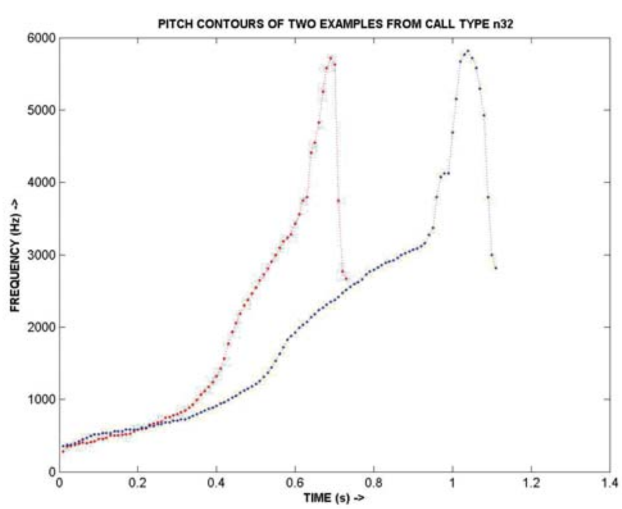

Fig. 3. Pitch contours of two examples from call type n32. The shorter contour is from the sound with spectrum in Fig. 2.

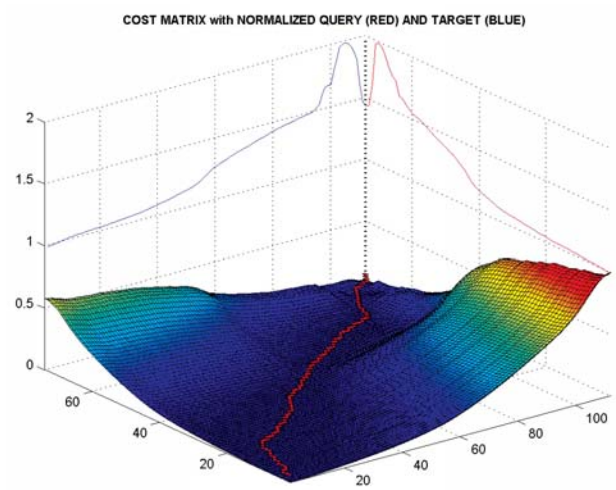

Fig. 4. Cost matrix with minimum cost path in bold red through the center. The shorter sound is called the query and the longer sound the target.

although the lengths of the calls will differ. For the automatic classification, a technique for quantitatively comparing curves of similar shape but different length is required. Dynamic time warping was widely used in the early days of speech recognition and more recently in musical information retrieval, and it is ideal for this task. The basic idea of DTW can be explained with an example using two sounds from the same group "n32."

A difference matrix is constructed from each number of sound 1 subtracted from each number of sound 2. This will give low values where the curves have similar values. From these numbers a cost matrix is constructed, which can be loosely thought of as a running sum of the differences between the two curves for all possible paths. The minimum path will follow the low numbers measuring overall differences in the best match of the two curves; this path can be traced and the final distance or dissimilarity is the last number attained in the minimum path. This can be visualized in Fig. 4 as the path of minimum effort through a mountainous terrain.

The "dissimilarity" or distance thus obtained is an excellent measure of contour differences. Identical signals will have a diagonal best path and a cost of zero (zero dissimilarity), while larger contour differences will have a correspondingly larger cost/dissimilarity. For classification these costs are a means of clustering the calls having the smallest dissimilarities.

Classification

The computer classification based on minimum dissimilarity within groups was compared to the human classification into call types for each of the two sets of whale sounds. For the Marineland calls using these distances and then running through the calculation a second time using the derivative of each of the contours (measuring the shape rather than the absolute value), an outstanding 99 % agreement with the human grouping was obtained. For the northern resident whales, dissimilarity matrices for the low frequency and high frequency calls were added to give an overall distance. Agreement with the perceptual classification was 90 % which is amazingly good given the similarity of contours in several of the sub-groups.

Conclusion

Dynamic time warping shows promise for automatic classification of killer whale call types. One of the most exciting applications of this technique would be the ability to monitor the movements and habitat-preferences of killer whales just by tracking sounds heard at remote monitoring stations. This will only be possible with systems developed to automatically process and identify calls heard at those locations so that the group producing them can be identified remotely.

Acknowledgment

JCB is very grateful to Dr. Eamonn Keogh for his MATLAB code which was adapted to give Fig. 3. Funding was provided to PJOM by Woods Hole Oceanographic Institution's Ocean Life Institute and a Royal Society fellowship.

Reference:

[1] H. Yurk, L. Barrett-Lennard, J. K. B. Ford, and C. O. Matkin, "Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska," Animal Behaviour **63**, 1103-1119 (2002).

Judy Brown is Professor of Physics Emerita at Wellesley College and Visiting Scientist at the Media Lab of the Massachusetts Institute of Technology. She is a Fellow of the Acoustical Society of America and has served on its Technical Committee on Musical Acoustics for over 10 years. Her interests lie in signal processing of musical sounds, and more recently classification of vocalizations of marine mammals.

Patrick Miller is a Senior Research Fellow at the Sea Mammal Research Unit in the University of Saint Andrews in Scotland, and recent holder of a Royal Society USA/Canada Fellowship. Miller completed the Ph.D. on vocal signals of killer whales in 2000 jointly from the Woods Hole Oceanographic Institution and the Massachusetts Institute of Technology. His research focuses on acoustic communication and social behavior of marine mammals, foraging and diving.

This article is based on paper 5aABb2 at the 151st ASA meeting in Providence.

At the College of Fellows luncheon at the Providence meeting Dr. Amar Bose gave an interesting lecture on a novel suspension system his company has developed for automobiles. Each of the four wheels is independently controlled by a computer which is modeled to react to potholes, bumps, and turns. A fascinating video showed cars leaning into turns (more like a motorcycle) rather than out. Also cars were shown smoothly going over speed bumps, which raised one obvious question from the audience (and maybe the police). The four wheel controllers can also be directly controlled, so that by appropriately raising and lowering the front and rear wheels the car was made to jump over a small hurdle—impressing everyone in attendance.

Dr. Amar Bose (center) is joined by Leo Beranek and Jan Weisenberger, Chair of College of Fellows. (Photo by Charles Schmid)