

TECHNICAL RESEARCH REPORT

Biologically-Inspired Acoustic Wear Analysis

by Varma S., Baras J.S., Shamma S.A.

CAAR TR 2001-1
(ISR TR 2001-11)



The Center for Auditory and Acoustic Research (CAAR) is a consortium of researchers from six universities working in partnership with Department of Defense laboratories and industry. CAAR is funded by the Office of Naval Research through a 1997 Department of Research Initiative.

Web site <http://www.isr.umd.edu/CAAR/>

BIOLOGICALLY-INSPIRED ACOUSTIC WEAR ANALYSIS

S. Varma, J.S. Baras and S.A. Shamma

Center for Auditory and Acoustics Research,
University of Maryland at College Park
varma@eng.umd.edu, baras@eng.umd.edu, sas@eng.umd.edu

ABSTRACT

We report on a novel method of acoustic wear analysis using spectral classification based on a model of mammalian audition. This approach uses biologically-inspired pre-processing filters that give a multi-resolution representation of the sound timbre. A Tree Structured Vector Quantizer (TSVQ) is used to create a classification tree based on the wear labels.

We have obtained encouraging results for tools of different diameters, cutting different materials. Trees trained on one kind of data seem to generalize well to new data sets.

1. INTRODUCTION

Much work has been done in real-time monitoring of machinery to detect faults as and when they occur, rather than wait until the next maintenance period. This way, unnecessary maintenance, as well as long runs in a faulty condition, can be avoided. In the case of a cutting tool, trying to cut with a blunt tool can lead to the breakage of the tool and degradation of the job, while pulling the tool off for frequent assessments are expensive in terms of the machinist's time. It is of interest to develop a method that can give an estimate of the wear from easily observable signals. This estimate can be used by the machinist to help her own intuition.

Experienced machinists can usually tell when a tool is too blunt to cut efficiently. Changes in cutting force or overheating are useful cues, but a more important indicator is the sound of the tool. Our goal in this work is to develop a system of classifying the acoustic signature of the tool according to the wear level.

This problem has many parallels with speech recognition, where the goal is vowel recognition or speaker identification. But there are some difficulties unique to the tool-monitoring problem. Classifying the sound is not the final goal here, as it is in the case of speech recognition. Here the aim is to classify sounds and then correlate these classes to the physical state of the tool. In addition, modifications are

needed because the training data are usually very sparsely labeled. For example, in our case, there are only two-three wear measurements for the whole lifetime of a tool. Having measurements at more frequent intervals is not possible in practice without disrupting the whole milling process.

Most previous work on estimating tool wear or damage from acoustic emissions has concentrated on using the power density spectrum in various ways; the simplest approach being just the average power of the sound signal [8], [12]. A more sophisticated way of using the power spectrum is to compare the total power in various sub-bands [3], [6]. These simple approaches give surprisingly good results in many cases. One approach which uses a learning expert system with torque and thrust information, in addition to vibration data is given in [5].

A rather more sophisticated approach is detailed in [9] where the author, working on the same data-set that we used, tries to isolate high-energy transients from the sound signal; one of the assumptions being that transients would be good indicators of chipping or fracture.

2. VIBRATION MODES OF COMPLEX MACHINERY

Determining the effect of wear on the acoustic emissions of a piece of machinery is complicated by the fact that machine tools have very complex vibration modes. Usually such machines can be modeled accurately only as 3-dimensional, non-linear, distributed systems. Non-linear phenomena like chatter are evidence for this [2]. Classical models of machine tool vibration have been almost exclusively linear, lumped-system approximations [13]. At least one author has applied statistical analysis in the form of an ARMA (Auto-regressive moving average) process driven by uncorrelated noise [10].

Since simple models of the tool surface - vibration relationship are not available, one is forced to look for non-parametric solutions to this problem. Our approach in this paper is to first extract a feature vector from the sound, and then do a non-model-based classification using Vector

Quantization [4]. Obviously, the performance of the classifier depends strongly on the proper selection of the feature vector.

We use filters based on a model of mammalian audition, which have been successfully used to classify vowel sounds [11]. This is followed by a tree structured classifier, based on vector quantization. Since one of our aims is to imitate what the human machinist does, this model was a natural choice.

3. AUDITORY FILTERS

We use two auditory filters, developed by Shamma et.al., for preprocessing. The first one is a model of the filtering and nonlinear operations that take place in the inner ear [15]. The second filter mimics the analysis of the filtered signal that take place in the primary auditory cortex [14].

3.1. INNER EAR

This filter (Fig 1) describes the mechanical and neural processing in the early stages of the auditory system. In the Analysis Stage, a bank of constant-Q filters, approximate the function of the eardrum and the basilar membrane in the cochlea with the continuous spatial axis of the cochlea as the scale parameter.

The Transduction Stage models the conversion of the mechanical displacements in the basilar membrane into electrical activity along a dense, topographically ordered array of auditory nerve fibers. This conversion can be well modeled by a three-stage process consisting of 1) a velocity coupling stage (time derivative), 2) an instantaneous non-linearity describing the opening and closing of the ionic channels and 3) a low-pass filter to describe the ionic leakage through the hair cell membranes.

The third stage called the Reduction Stage effectively computes an estimate of the spectrum of the stimulus, through a *lateral inhibitory network* (LIN). The details can be found in [15]. The output of this filter is a spectral estimate of the input.

3.2. THE AUDITORY CORTEX

The second filter that acts on the output of the first filter is based on the action of the primary auditory cortex (A1). In the A1, the 1-D acoustic spectrum is analyzed along three feature axes: the *spectral symmetry* on the ϕ axis, the *local bandwidth* on the scale s axis, and the *frequency components* on the tonotopic x axis. This analysis can be thought of as a local affine wavelet transform of the acoustic spectrum. The scale axis s gives a multi-scale representation of the spectrum. The lower the resolution, the broader the bandwidth that the cell is attuned to.

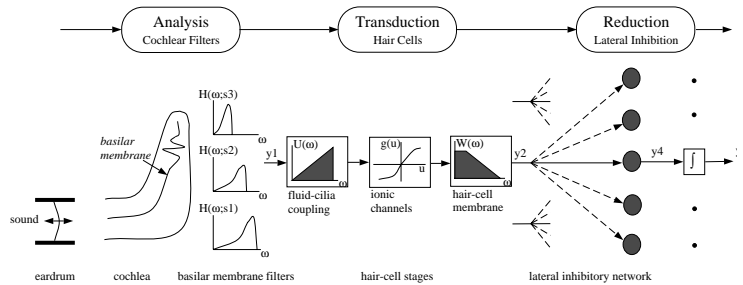


Figure 1: Spectral processing of sound stimuli in the inner ear

This approach is similar to other methods like discrete cosine transform (DCT) or cepstrum based frameworks that are often employed for spectral shape analysis. However cepstral methods differ from the cortical approach in that the cortical approach is local along the tonotopic axis x , where the cepstral coefficients capture global features of the spectral shape.

4. TREE STRUCTURED VECTOR QUANTIZER (TSVQ)

TSVQ is an example of a classification tree where the given test vectors are classified stage by stage, with each stage giving a sharper classification than the previous. Each node of the tree is associated with a centroid, which can be thought of as a paradigm for a particular class (Fig 2). All test vectors start out by belonging to the root node. Then the vector is compared with the centroids of all nodes which are children of the node it currently belongs to. The vector is classified into the child with the centroid that is “closest” to it according to some metric. The vector eventually ends up in a leaf node, and is assigned a class according to the class of the leaf node.

Making a vector quantizer (VQ) in the form of a tree offers several advantages. Firstly, if the tree is more or less balanced, the number of comparisons that have to be made are $O(\log n)$ where n is the total number of partitions of the vector space. For a VQ on one level, we would have to make $O(n)$ comparisons. This improvement in search efficiency can be a big factor when we have a large number of classes (as in the case of classifying radar returns [1]). Also, as explained later, we can use parallel TSVQ techniques to further reduce the search time.

Secondly, the way the underlying vector space is split at each node is frequently indicative of natural partitions in the data-set. The sequence of features that play the most important role in the partitioning at each level give a natural

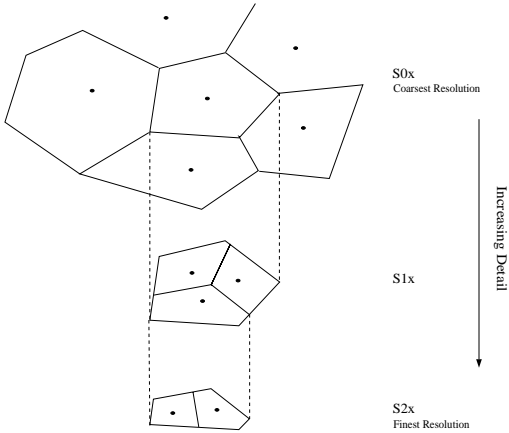


Figure 2: Multi-resolution tree structured VQ

ordering of their importance in the classification. Thus in the case of the radar return classification mentioned above, the tree structure gives a tessellation similar to the *aspect graph* in computing literature.

4.1. MULTI-RESOLUTION TSVQ (MRTSVQ)

One special kind of VQ classifying tree is the Multi-Resolution TSVQ (MRTSVQ). Any particular test vector in an MRTSVQ is represented in multiple resolutions or scales. One obvious method of creating such a representation is through affine wavelet transforms. The auditory cortex filter we use is an example of a biological multi-resolution transform that follows from studies of the primary cortex.

Mathematically, for a given vector x we create a multi-resolution representation

$$S_0x, S_1x, S_2x, \dots, S_Jx \quad (1)$$

Where $S_i x$ is the vector x at the i^{th} resolution in terms of some multi-resolution transform S_i . At each level i in the tree, the i^{th} resolution vector $S_i x$ is compared against the centroids of the nodes in that particular level of the sub-tree. The centroids at a particular level are also represented at the corresponding resolution. The vector is classified into the node that has the nearest neighbor centroid. At the next level, the next higher resolution of the vector is used for the comparison.

This method of classification offers one advantage over the unembellished TSVQ. At the higher levels, where more comparisons have to be made, we can use a vector with lesser number of bits, thus doing many simple computations. As we go down the tree (and sharpen our classification), we do lesser number of progressively longer distance-

calculations. Small amounts of information, in the form of a coarse approximation of a suitable representation of the signal, are used first to provide partial classification. Progressively finer details (more information) are added until satisfactory performance is obtained. This computational advantage is very important in online algorithms.

5. TRAINING

In our way of combining class labels in growing TSVQ, we build a tree for each class, using only the appropriately labeled data. This method, usually called Parallel TSVQ, gives better results than making one tree for all the classes combined. In the combined tree, an initial wrong misclassification into one particular sub-tree can end in a vector being incorrectly classified. This problem is avoided, to a great extent, in the parallel case. The Parallel TSVQ is also quicker to execute when we have a large number of classes. Testing on each tree can be done in parallel, which reduces computational time.

5.1. PREPROCESSING OF TRAINING DATA

The sound data is cut up into frames, each frame corresponding to the sound in one revolution. Each frame is then cut into four sub-frames, corresponding to each quarter (flute) of the tool. The sub-frames are then passed through the inner ear and auditory cortex filter to obtain a set of multi-resolution vectors describing the timbre spectrum of the sound. We use an average of the four sub-frames in each frame as input to the tree growing algorithm.

5.2. TREE GROWING ALGORITHM

We use a tree-growing algorithm that has been used by Baras and Wolk [1] for classifying radar returns. The details of the algorithm can be found in the above reference. Essentially, the tree algorithm uses the Linde-Buzo-Gray (LBG) algorithm [7] for VQ at each level of the tree. At each level, the tree algorithm starts with an initial fixed number of centroids. The LBG algorithm is used to find a distribution of centroids that correspond to a local minimum in the expected squared error distortion. Then an additional centroid is introduced and LBG applied again to find the expected distortion. If the change in distortion from the addition of this new centroid is greater than a fixed fraction of the original distortion, another centroid is introduced and the process repeated.

If the change in distortion is lesser than then fixed fraction, the algorithm goes to the next level. The cell in the current level is fixed and the leaf node with the highest value of the distortion is then split at the next lower level (higher

resolution). This process goes on until a stopping criterion is satisfied. The stopping criterion we used was the final number of leaf nodes.

6. TESTING

Testing was done on data belonging to different tool diameters and job material. Tools that had not been used in training were used in the testing procedure. The preprocessing was similar to what was done for training. Each vector was dropped down all trees and the distance to the centroids of the leaf nodes it fell into, was compared. The vector is assigned a wear-class according to the wear level of the tree that gives the least distance from the centroid to the vector.

This way, we get a time series of wear-class prediction for all the frames for all the passes. Next we take a sliding window of 500 frames and find the mean wear estimate for this window. Plots of the mean wear estimate vs. tool-life in frames (revolutions) is shown in Fig 3, 4 and 5.

For the case of a 0.5" dia. tool cutting 4340 steel, Fig 3, it is apparent that our method has picked up features in the sound that seem to be correlated to the tool-life and the wear of the tool. The periodic variation in the wear estimate is a result of the different passes. The actual sound made by the tool is not just a function of the state of the tool, but also depends on the position on the job the tool is presently at. The character of the sound at the beginning and the end of the cut is very different from that in the middle, with all other conditions remaining the same. The wear estimate at the start and end of the pass are higher than that in the middle. This could correspond to the observation that the *wear rate* at the starting and ending of the job is higher than in the middle. This might point to the fact that the sound of the tool is more properly classified into a particular *wear and wear rate combination*, and not just wear.

Fig 4 shows the results of testing data corresponding to a 0.5" dia. tool cutting a titanium job on the same classification tree. Here also we see the gradual increase in the tool wear, though the way it increases is different from that in Case 2. Fig 5 shows the results of the classification on data for a 1.0" dia. tool cutting 4340 steel.

7. CONCLUSIONS

The mammalian ear model coupled with a TSVQ seems to pick out features that are strongly indicative of wear. Furthermore, trees trained on one particular tool-job configuration seem to generalize easily to other configurations. This indicates that our features are not tool or material specific, but are characteristic of the cutting process in general. Algorithms that seek to predict tool wear based on just the spectrum cannot do this.

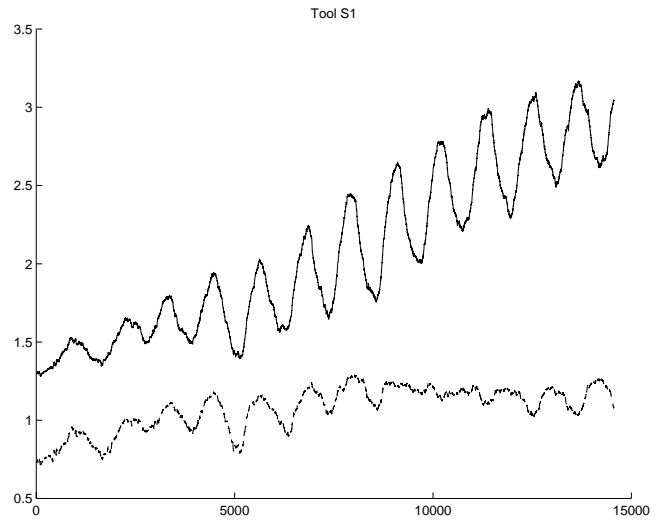


Figure 3: Mean and variance of wear estimates versus time (in frames) for tool s1

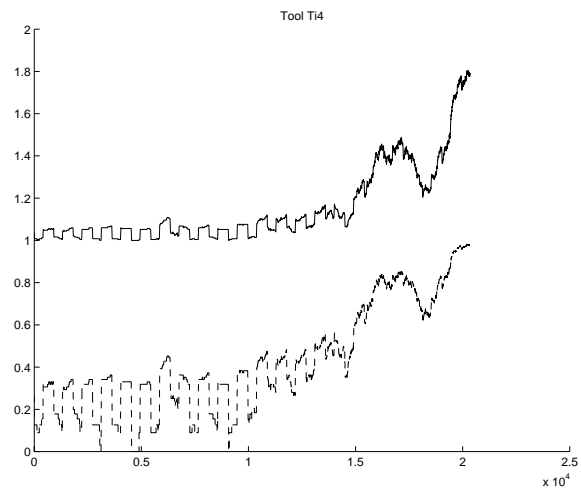


Figure 4: Mean and variance of wear estimates versus time (in frames) for tool ti4

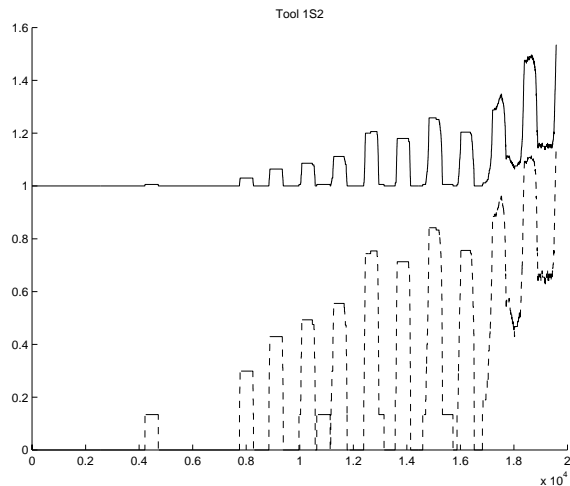


Figure 5: Mean and variance of wear estimates versus time (in frames) for tool 1s2

Two directions for future research are

1. Developing a model of the tool wear process in terms of dynamical systems like Hidden Markov Models.
2. Investigating whether classification of the sound into a combined (wear, wear-rate) state rather than just a wear state gives any additional benefit. Physically, it has been noticed that sounds that indicate chatter and other nonlinear phenomena correspond to a particular wear-rate, rather than a particular wear. The difficulty in this approach would be a lack of frequent measurements.

Acknowledgements

The authors would like to thank Dr. Gary Bernard of Boeing Co. for giving us access to the machining data. This research was funded in part by grant N-000149710501EE from the Office of Naval Research.

8. REFERENCES

- [1] Baras, J.S., and Wolk, S.I. *Hierarchical wavelet representations of ship radar returns* Institute for Systems Research Technical Report-TR 93-100, 1993, University of Maryland at College Park.
- [2] Berger, B.S., Rokin, M. and Minis, I. *Complex dynamics in metal cutting* Quarterly of App. Math., Vol. 51, 1993, pp 601-612
- [3] Chen, M-F, Natsuake, M. and Ito, Y. *Preventive monitoring system for main spindle of machine tool - comparison of signal processing methods* Proc. 27th Inter. MATADOR Conf., April 1988, pp. 179-186
- [4] Gersho, A. and Gray, R.M. *Vector quantization and signal compression* Kluwer Academic Press, 1991
- [5] Hong, S.Y. *Self-learning system for knowledge based diagnosis of drill condition in fully automatic manufacturing system* Proc. SPIE, Vol. 1707, 1992, pp. 195-206
- [6] Inasaki, I. and Yonetsu, S. *In-process detection of cutting tool damage by acoustic emission measurement* Proc. 22nd Inter. Machine Tool Design and Res. Conf., 1981, pp. 261-268
- [7] Linde Y., Buzo, A. and Gray, R.M. *An Algorithm for Vector Quantizer Design* IEEE Trans. on Comm., Vol. COM-28, No. 1, Jan. 1980, pp. 84-95
- [8] Noh, A. and Kobayashi, A. *Cutting tool failure detector* "Adv. in Hard Material Tool Tech." Proc. Inter. Conf. on Hard Materials Tool Tech., 1976, pp. 414-426
- [9] Owsley, L.M.D., Atlas, L.E. and Bernard, G.D. *Self-Organizing Feature Maps and Hidden Markov Models for Machine-Tool Monitoring* IEEE Trans. on Sig. Proc., Vol. 45, No. 11, Nov. 1997, pp. 2787-2798
- [10] Pandit, S.M., Subramanian, T.L., and Wu, S.M. *Modeling machine tool chatter by time series* ASME J. of Eng. for Industry, 1975, pp 211-215
- [11] Shamma, S.A. *The acoustic features of speech phonemes in a model of the auditory system: Vowels and unvoiced fricatives* J. of Phonetics, Vol. 16, 1988, pp. 77-91
- [12] Shumsheruddin, S. and Lawrence, J.C. *In-process prediction of milling tool wear* Proc. 24th Inter. Machine Tool Design and Res. Conf., 1983, pp. 201-214
- [13] Tobias, S.A. *Machine-tool vibration* J. Wiley, 1965
- [14] Wang, K. and Shamma, S.A. *Spectral shape analysis in the central auditory system* IEEE Trans. on Speech and Audio Proc., Vol. 3, No. 5, Sept. 1995, pp. 382-395
- [15] Yang, X., Wang, K. and Shamma, S.A. *Auditory representations of acoustic signals* IEEE Trans. Info. Theory, Vol. 38, No. 2, March 1992, pp. 824-839