# DESCRIPTION OF THE MULTIPLE LOOK APPROACH FOR CALCULATING UNSTEADY LOUDNESS

**Helen Ule[1], and Colin Novak[2]**
[1]Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4 ule@uwindsor.ca
[2]Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4 novak1@uwindsor.ca

## 1. INTRODUCTION

Human perception of the quality of sound of a product or noise source is important because it is tied to human comfort. Therefore, it is important for design engineers to design various industrial products related to automotive, consumer electronics, computer, etc. so that sound quality is enhanced. The ongoing work presented in this paper is focused on the development and improvement of the specific psychoacoustic metric of loudness such that it can better correlate with real human perception. This metric characterizes the strength attribute of the human auditory response in terms of the perception of sounds from quiet to loud and has become a very important sound quality metric for the acoustic evaluation of product noise. However, no calculation method for determining the loudness of sounds which vary with time has been successfully correlated with all experiments involving human perception. The fundamental goal of this research is to develop an approach which will calculate the time varying loudness using the multiple-look model which is thought to be able to better process human perception characteristics not identified in other proposed models.

## 2. CALCULATING UNSTEADY LOUDNESS

For the calculation of unsteady loudness, two methods exist. The first is referred to as a Zwicker base approach since it is an extension of the stationary calculation method developed by Eberhart Zwicker (1) and provides the foundation for the loudness calculation method given by ISO 532 (1975). This method has been expanded upon and provides the basis for the draft DIN 45631/A1 for calculating unsteady loudness. A second approach has been developed by Moore and Glasberg (2) which uses a method which they called long term integration.

For the Zwicker based approach, the loudness is based on the distribution of the specific loudness along the critical band scale only here they are treated as time dependant values. For each of the critical bands, the effect of temporal masking is accounted for by applying corrections based on post-masking data developed by Zwicker. It has been found that in the presence of impulsive signals, the apparent loudness decreases with the duration of the impulse. This is accounted for in the Zwicker approach. Finally, a network summing all specific loudness values along the critical bands is applied in order to get the overall loudness value.

Glasberg and Moore's temporal integration approach involves combining information over time which they say is to improve detection or discrimination over the Zwicker approach and is done in 5 steps. First, the sound is sampled at a rate of 32 kHz and put through filter representing the function of the outer and inner ear. Next, the excitation pattern, or specific loudness, is calculated from a fast fourier transform of the filter signal. The specific loudness is then integrated to get the instantaneous loudness. Finally it is averaged to get the short term and long term time varying loudness.

## 3. JUSTIFICATION OF MULTIPLE LOOK APPROACH

The classical approach to the calculation of time varying loudness is temporal integration, often called temporal summation, which involves the combining of information over time to improve detection and is often thought of as a simple accumulation process such as energy integration (2). However, recent studies (3,4,5) suggest there is a better approach to the calculation of time varying loudness which is to consider the combination of audible information from multiple independent looks (2).

Viemeister and Wakefield (4) studied the effect of signal duration on detection and discrimination thresholds of sound pressure levels. From experiments, they have shown that pulse tones with temporal separations greater than 10 ms are processed independently which subsequently draw a conclusion that the classical view of long-term integration is not an acceptable method for the characterization of signals containing gaps which are often prevalent in a non-stationary noise signal. Instead, the results are consistent with the idea that the input signal is sampled at a fairly high rate and each sample, or "look" is stored in memory for selective access and processing. It is this idea that provides the foundation for the multiple-look model as a more accurate perceptual representation, particularly in the presence of temporal gaps.

Moore (2) has performed various experiments which demonstrate that, "Information extracted from one part of a sound may influence the evaluation and interpretation of information extracted from another part, at a different time." These results not only support the conclusions of Viemeister and Wakefield, but also do so through more rigorous experimentation. The outcome of this work suggests that the temporal integration which occurs for unsteady sounds cannot be characterized by a basic summation or accumulation process, but instead supports the approach of the multiple-look concept.

The work done by Pedersen (16) on the various stages of analyzing time varying sounds is also of importance. Pedersen was able to demonstrate that different types of temporal processing are involved in hearing, each of which are dependent on the specific task given to the listener. The result of this work reinforces the reasoning as to why long-term integration techniques do not correlate with the experiments of Viemeister and Wakefield. More importantly, this work yields further justification to seek alternative approaches such as the multiple-look model proposed in this research plan.

While algorithms and computer programs have been proposed for time-varying loudness, a version which utilizes the multiple-look concept has not been realized. It is anticipated that the creation of such an algorithm would provide the ability to determine the loudness of an unsteady signal which would include the ability to discriminate varying thresholds due to short-term and burst temporal effects.

## 4. METHODOLOGY OF MULTIPLE LOOK APPROACH

To overcome the shortcomings of the present unsteady loudness calculation techniques, the multiple look method is based on a short time constant process of 2 to 3 ms where decisions are based upon selected "looks" and can in principle account for the temporal resolution for modulation detection, gap detection, etc. Another important point of the multiple-look model is that it allows for intelligent processing of the "looks". As these are stored in memory, they can be selectively used for further processing and decision making.

While it is believed that the multiple look approach will overcome the issues that the other methods have, the present research is focused on the application for gap detection. The following is the proposed approach:

The time signal is first aacquired in short step durations of 2 ms for which an FFT is applied for each stepped duration,as well as for longer, or multiple step durations, to account for low frequency components of the sound. Next the sampled signal has been put through the response filters representing the inner and outer ear functions. From here, the specific loudness of the signal can be calculated for each of the small duration segments, or "looks", which are sampled at a high rate. The program then stores the vector of "looks" in temporary memory and is put through standardized computations and comparisons. This is analogous to how non-varying sounds are compared to values in "look-up" tables for the calculation of stationary loudness in the present ISO 532B method, only here, it is done for extremely short duration signals in very quick succession. The treatment of the "looks" is dependent on the task, and for this study, gap detection is targeted. As such, using the 2 ms loudness, it must be decided if a gap is present by looking at three successive samples for immediate drops in level. If no gap is detected, the process returns to the beginning of the signal and repeats for up to 100 ms and apply long term integration unless a gap is detected before 100 ms. If a gap is found to exists, the program modifies the threshold detection for loudness of the 3 sampled steps of data. Finally with this long term integration is applied to the eventual 100 ms of signal.

## 5. CONCLUSSION

This paper described an alternative approach to calculating time varying loudness using a multiple look model. This approach involves the dissection of a time signal into small duration segments which are sampled at a very high rate. These multiple looks are then processed independently and are thought to be able to better process the perception characteristics, specifically gap detection, not identified in the temporal integration model traditionally used. While this work is still in progress, the algorithm being developed is hoped to proven to result in the calculation of loudness with better correlation to attributes of human perception.

## 6. REFERENCES


1. Zwicker, E., Erratum: "Procedure for calculating loudness of temporally variable sounds". *Journal of the Acoustical Society of America*, 63(1), 283.
2. Moore, B. (2003). Temporal integration and context effects in hearing. *Journal of Phonetics*, 31, 563-574.
3. Glasberg, B., Moore, B., (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5), 331-342.
4. Viemeister, N., Wakefield, G. (1991). Temporal integration and multiple looks. *Journal of the Acoustical Society of America*, 90(2), 858-865.
5. Pedersen, B. (2006). *Auditory Temporal Resolution and Integration – Stages of Analyzing Time-Varying Sounds*. Doctoral dissertation, Department of Acoustics, Aalborg University.
6. Pedersen, B., Ellermeier, W. (2005). Temporal and spectral interaction in loudness perception. *Journal of the Acoustical Society of America,* 117(4), 2397-2407.