ICSV14
Cairns • Australia
9-12 July, 2007

14th International Congress on Sound & Vibration

FFT BASED COMPLEX CRITICAL BAND FILTER BANK AND TIME-VARYING LOUDNESS, FLUCTUATION STRENGTH AND ROUGHNESS

HE Lingsong[1], Malcolm J. Crocker[2] and Zhou Ran[2]

[1]School of Mechanical Engineering, Huazhong University of Science and Technology
Wuhan, Hubei, 430074, China
[2]Department of Mechanical Engineering, Auburn University
201 Ross Hall, Auburn University, AL 36849-5341, USA
mcrocker@eng.auburn.edu

Abstract

This paper describes a new algorithm for producing a complex critical band filter bank and its application in computing time-varying loudness, fluctuation strength and roughness. A digital Butterworth filter bank is used to define the frequency characteristics of the critical band filter bank and then transform it into a complex critical band filter bank by use of the Hilbert transform. An FFT transform is used to transform the noise signal into the frequency domain, and then the noise signal is separated into critical bands by multiplying it by each critical band filter. Finally multi-IFFT transforms are employed to transform each filtered signal into the time domain in two orthogonal parts or into a complex signal form. Thus the envelope curve can be obtained for each bark band from the amplitude of the complex signal. The instantaneous energy of the envelope curve for each critical band is the instantaneous loudness in that critical band and the depth of the envelope curve of each critical band is proportional to the fluctuation strength and roughness in that critical band.

1. INTRODUCTION

The Zwicker method has come to be seen as the most useful and the most commonly used approach to calculate the loudness needed in sound quality software. In the Zwicker method, one-third octave sound pressure levels are used as input data and are transformed into the loudness of critical bands by a psychoacoustic model; then the loudness of critical bands are summed to obtain the total loudness. In most cases, the Fast Fourier transform (FFT) is employed to calculate the power spectrum and the one-third octave bands of the signal, then to calculate the loudness. For the stationary loudness analysis, there is no problem. For the time varying loudness analysis, a problem arises. The FFT has a fixed spectral resolution, which depends on the length of the signal time window. With the signal time window of 100 milliseconds, the spectral resolution is 10 Hz. A signal with a time window of 10 milliseconds has a spectral resolution of 100 Hz. When the spectral resolution of the signal is low, the one-third octave resolution of the signal is low also. The loudness calculated is incorrect as well.

In the definitions of roughness and fluctuation strength, the time varying loudness is the basic parameter used to obtain these two psychoacoustic quantities. For the calculation of the roughness from the time varying loudness, a 3 milliseconds time resolution is a minimum requirement. This means that a 333 Hz spectral resolution is used in calculating the loudness by the FFT. In this paper, a new FFT-based filter bank algorithm is presented to calculate the time varying critical band spectrum, which is the key to obtain the time varying loudness. In addition some new ideas are outlined for the use of the time varying loudness to determine the fluctuation strength and roughness.

2. ESTIMATIONS OF THE TIME VARYING LOUDNESS

2.1 Critical bands

In the Zwicker method, the audible frequency range is divided into 24 critical bands. Table 1 shows parts of the arrangement [1]. The bands are similar to the frequency bandwidths of the one-third octave bands. The one-third octave bands are replaced by critical band spectra in the calculation of the loudness in the Zwicker method.

Table 1. Critical Bands

Critical Band z	1	2	3	4	5	6	7	8	…
Low Frequency (Hz)	50	150	250	350	450	570	700	840	…
High Frequency (Hz)	150	250	350	450	570	700	840	1000	…
Central Frequency (Hz)	100	200	300	400	510	630	770	920	…

2.2 Critical Band Filter Bank

The critical band filter bank is a group of filters. Each filter corresponds to a critical band. At any frequency, the sum of the squares of each filter's amplitude response is equal to one. This means that no energy losses occur in the filtered signals.

$$\sum_{b=1}^{24} \left| H_b(\omega) \right|^2 = 1 \quad (1)$$

where $| H_b (\omega) |$ is the amplitude response of the filter in the critical band b.

To design filters that satisfy these requirements, a band-pass Butterworth filter is constructed to define the critical band filter.

$$\left| H_b(\omega) \right|^2 = \frac{(\frac{\omega}{\omega_{b2}})^{2p}}{(1+(\frac{\omega}{\omega_{b1}})^{2p})(1+(\frac{\omega}{\omega_{b2}})^{2p})} \qquad b = 1,2,...,24 \quad (2)$$

where p is the order of Butterworth filter, ω_{b1} is the lower cut-off frequency, and ω_{b2} is the upper cut-off frequency of the critical band filter in the band b. In order to obtain an acceptable digital filter, the magnitude response of the analog filter in Eq. (2) is limited to the Nyquist range {-π< ω < π } in the bilinear transform.

$$|H_b(\omega)|^2 = \frac{\left(\frac{tg(\omega/2)}{tg(\omega_{b2}/2)}\right)^{2p}}{\left(1+\left(\frac{tg(\omega/2)}{tg(\omega_{b1}/2)}\right)^{2p}\right)\left(1+\left(\frac{tg(\omega/2)}{tg(\omega_{b2}/2)}\right)^{2p}\right)} \quad -\pi<\omega<\pi \tag{3}$$

When a sampling frequency H_s is fixed, the normalized cut-off frequency of the critical band filter can be obtained from the data in Table 1.

$$\omega_{b1} = 2\pi * \frac{f_{b1}}{F_s}, \omega_{b2} = 2\pi * \frac{f_{b2}}{F_s} \quad b = 1,2,...,24 \tag{4}$$

where f_{b1} is the lower cut-off frequency, and f_{b2} is the upper cut-off frequency of the critical band filter in the critical band *b*. Fig. 1 shows the magnitude responses of a Butterworth filter of different order *p*.

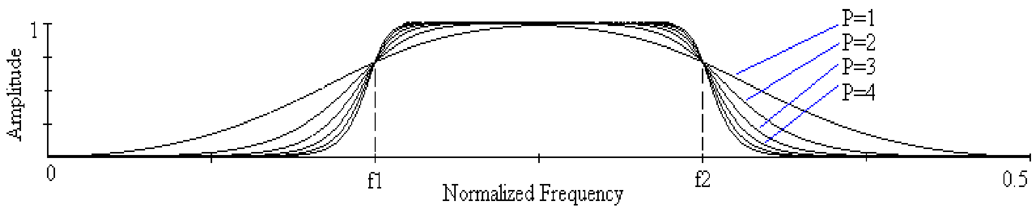

Figure.1 Magnitude responses of a Butterworth filter in different order *p*

By obtaining the cut-off frequency for each critical band filter from Eq. (4) and substituting it into Eq. (3), critical band filters can be built and a critical band filter bank can be created.

2.3 Orthogonal Complex Filter and Envelope Filter

To determine the time varying loudness, it is much more important to obtain the instantaneous intensity of the signal in each critical band, than the signal filtered in each band. To obtain the instantaneous energy of the signal, an orthogonal complex filter can be used to filter the signal and to obtain the envelope of the signal, which is proportional to the energy of the signal.
An orthogonal complex filter is composed of a real filter and an imaginary filter. Both filters have the same magnitude response but have a 90^0 phase difference between them.

$$H_b(\omega) = H_{br}(\omega) + jH_{bi}(\omega) \tag{5}$$

A simple method is to define $H_{br}(\omega)$ as a symmetrical filter and to define $H_{bi}(\omega)$ as an anti-symmetrical filter.

$$H_{br}(\omega) = \begin{cases} |H_b(\omega)| + j0 & ,0 \le \omega \\ |H_b(\omega)| + j0 & ,0 > \omega \end{cases}, H_{bi}(\omega) = \begin{cases} 0 + j|H_b(\omega)| & ,0 \le \omega \\ 0 - j|H_b(\omega)| & ,0 > \omega \end{cases} \tag{6}$$

This leads to

$$H_b(\omega) = H_{br}(\omega) + jH_{bi}(\omega) = \begin{cases} 0 + j0 & ,0 \le \omega \\ 2|H_b(\omega)| + j0 & ,0 > \omega \end{cases} \quad (7)$$

This takes the form of the Hilbert transform. When the signal is filtered with an orthogonal complex filter, the real part and imaginary parts of the filtered signals have a 90^0 phase difference and the magnitude of the filtered signal is equal to the envelope of the filtered signal.

$$\begin{aligned} h_b(t) &= h_{br}(t) + j * h_{bi}(t) \\ y_b(t) &= y_{br}(t) + jy_{bi}(t) = x(t) * h_b(t) = x(t) * h_{br}(t) + jx(t) * h_{bi}(t) \end{aligned} \quad (8)$$

where $h_b(t)$ is the impulse response of the complex filter $H_b(\omega)$,and $y_b(t)$ is the complex filtered signal of the original signal $x(t)$. The envelope of the filtered signal is

$$z_b(t) = \sqrt{y_{br}{}^2(t) + y_{bi}{}^2(t)} \quad (9)$$

Fig. 2 shows the envelope of a "beat" wave. Compared with the original signal, its envelope is usually a low frequency signal. This means that less sampled data are needed, and the processing is easier when the original audio signal is separated into twenty-four filtered signals by the critical band filter bank.

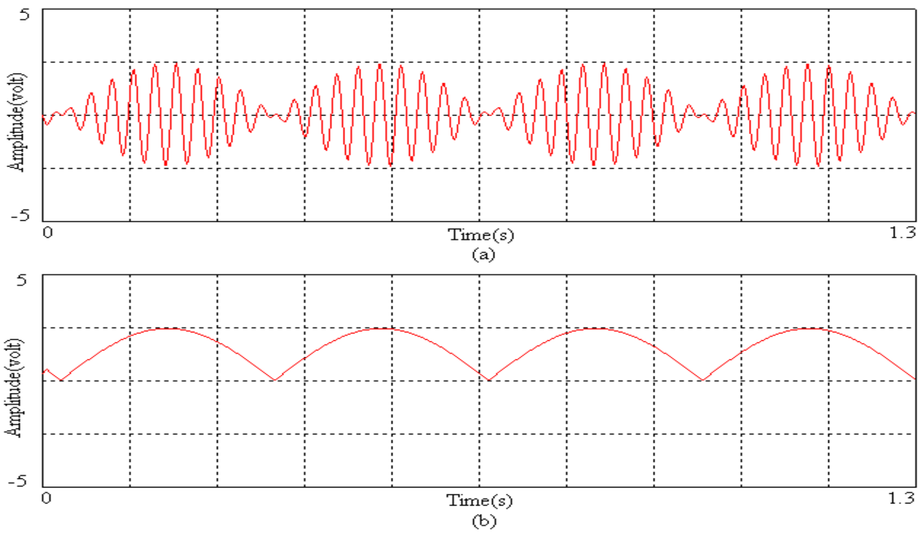

Figure.2 (a) The "beat" wave and (b) its envelope

2.4 FFT Based Envelope Filter

The signal filtering process can be viewed as the inner product of the signal and the filter in the time domain, and the process can also be carried out by calculating the product of the signal and the filter in the frequency domain.

$$\begin{aligned} Y_b(\omega) &= X(\omega) \times H_b(\omega) \\ y_b(t) &= F^{-1}[Y_b(\omega)] \\ z_b(t) &= \sqrt{y_{br}{}^2(t) + y_{bi}{}^2(t)} \end{aligned} \quad (10)$$

where $X(\omega)$ is the Fourier transform of the signal $x(t)$, and $H_b(\omega)$ is the critical band filter defined in Eq. (7). Fig. 3 shows the envelope filter algorithm obtained using the FFT.

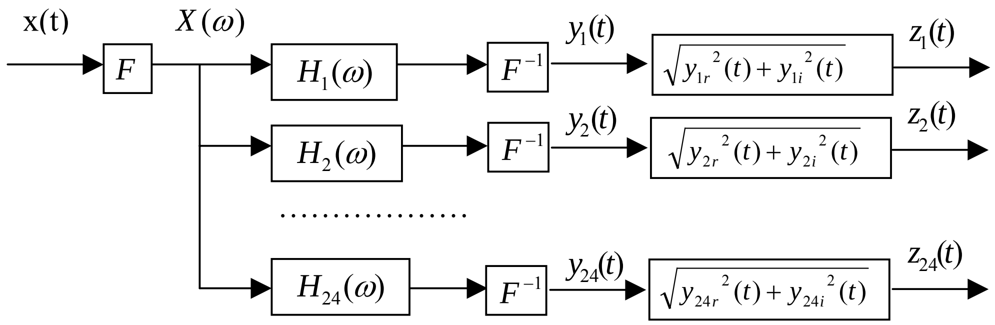

Figure. 3 Envelope filter algorithm using FFT

2.5 Short time energy of critical bands

The short time energy of each critical band is defined as

$$E_T(n) = \frac{0.5}{T}\int_0^T z_b^{\,2}(t) \tag{11}$$

where T is the length of time window. From Eq. (9), the short time energy of the envelope of the signal is twice as large as the short time energy of the signal, so a factor of 0.5 is introduced into the formula.

Four time windows are used for different purposes. A 2.5 milliseconds time window is used for the instantaneous loudness. A 20 milliseconds time window is used for loudness when it is varying very quickly. A 125 milliseconds time window is used for loudness varying quickly. A 1000 milliseconds time window is used for loudness varying slowly. Fig. 4 shows a white noise signal and its short time energy in critical bands. The length of the time window is 125 milliseconds.

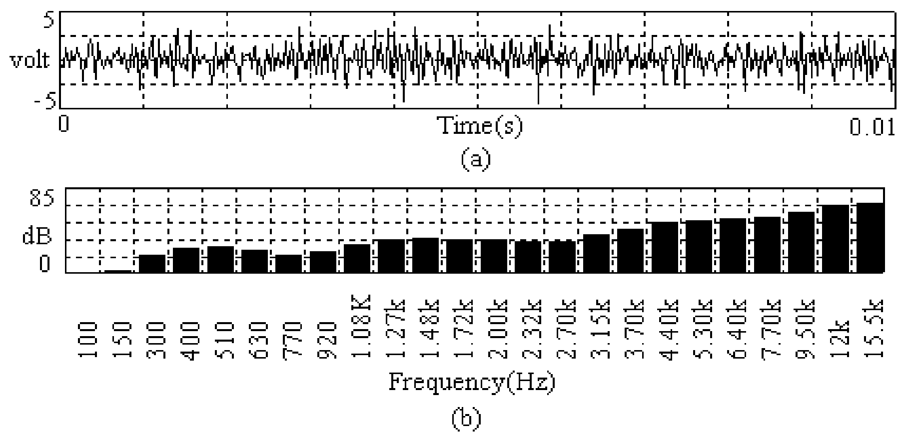

Figure. 4 (a) White noise signal and (b) its critical band short time energy

2.6 Time-varying Loudness

Two methods are used to obtain the time-varying loudness. The first method uses the critical band short time energy to replace the one-third octave band in the Zwicker method. It is much the same as the normal Zwicker method. Fig. 5 shows the critical band short time energy and the Zwicker critical band response.

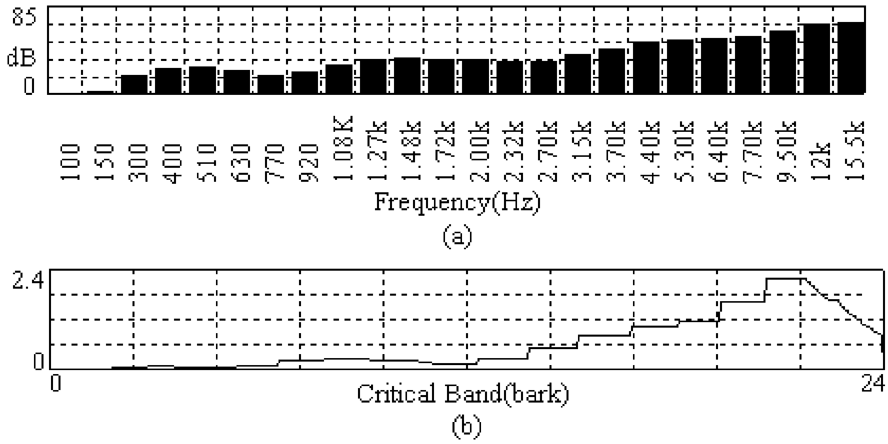

Figure. 5 (a) Critical band short time energy and (b) Zwicker critical band response

The second method uses the equal sound pressure level contour that is obtained from the equal loudness level contour to transform the critical band short time energy into the critical band loudness [3], [4]. The equal loudness level and equal sound pressure level contours are shown in Fig. 6.

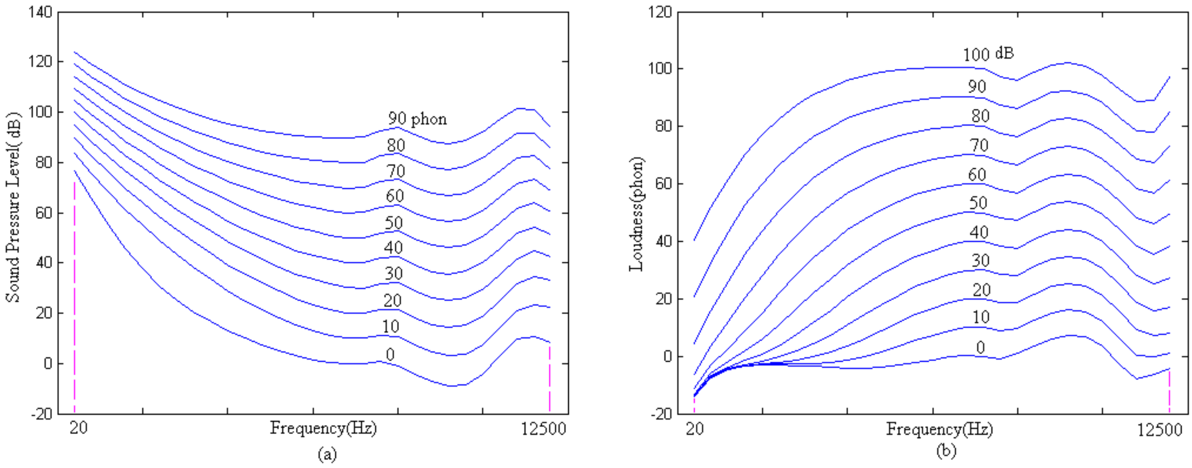

Figure. 6 (a) Equal loudness level contour and (b) equal sound pressure level contour

The time varying total loudness can be calculated from the time varying critical band loudness using the ISO R 532 A (Steven's method) [2].

$$N_{Total} = (1-0.15)Max(N_b) + 0.15\sum_{b=1}^{24} N_b \quad (24)$$

where N_{Total} is the time varying total loudness, and N_b is the time varying critical band loudness in the critical band *b*.

Fig. 7 shows a time varying loudness analysis result of a Ford Mustang sports car. Fig. 8 shows a time varying loudness analysis result of a Jaguar S-type luxury saloon car. One can see that the road noise of the sports car fluctuates considerably in time, while the road noise of the Jaguar car is much steadier. The results are the similar to the subjective assessment of the noise made by the human subjects.

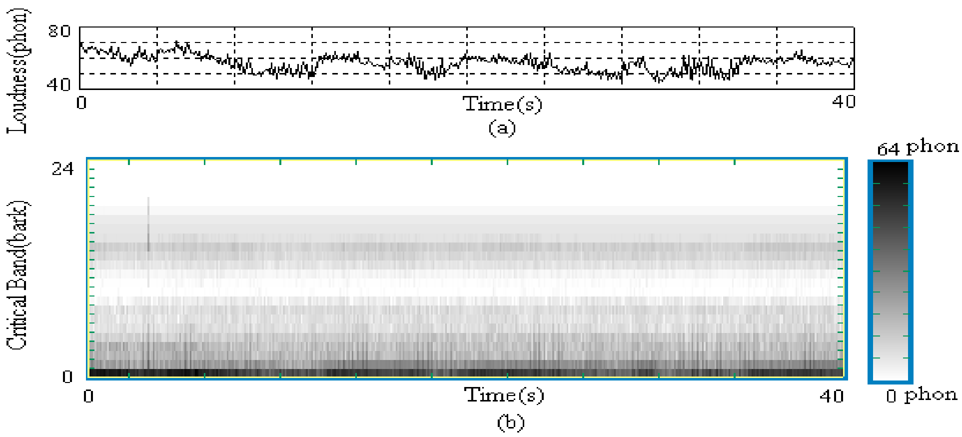

Figure. 7 (a) Time varying loudness and (b) time varying critical band loudness of the Mustang sports car

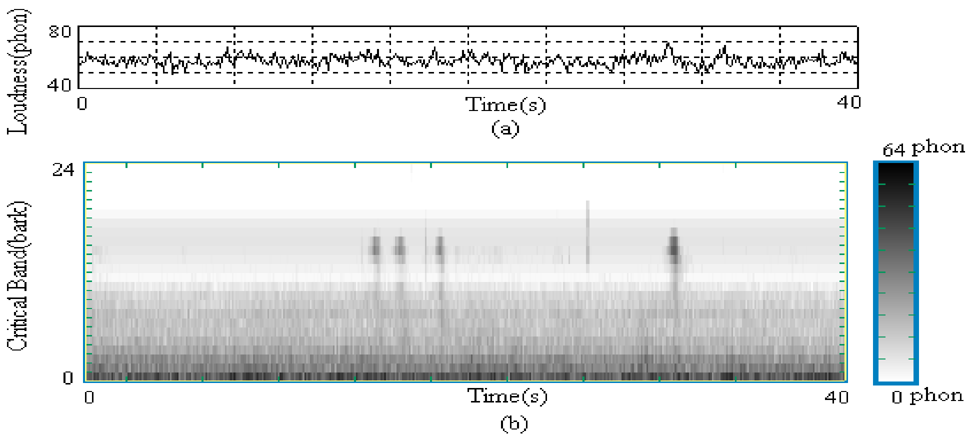

Figure. 8 (a) Time varying loudness and (b) time varying critical band loudness of the Jaguar saloon car

2.7 Fluctuation Strength and Roughness

The time varying total loudness shows the variation in loudness. It can be used to estimate the fluctuation strength and the roughness. Figure 9 shows a white noise signal that is modulated by a 4 Hz signal at a depth of 80%. Fig. 10 shows a white noise signal that is modulated by a 20 Hz signal at a depth of 80%.

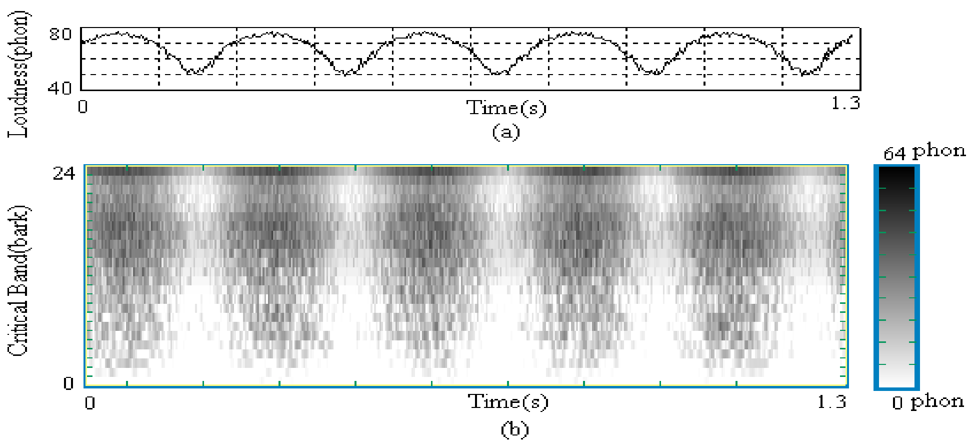

Figure. 9 (a) Time varying loudness and (b) time varying critical band loudness of a 4 Hz modulated white noise signal

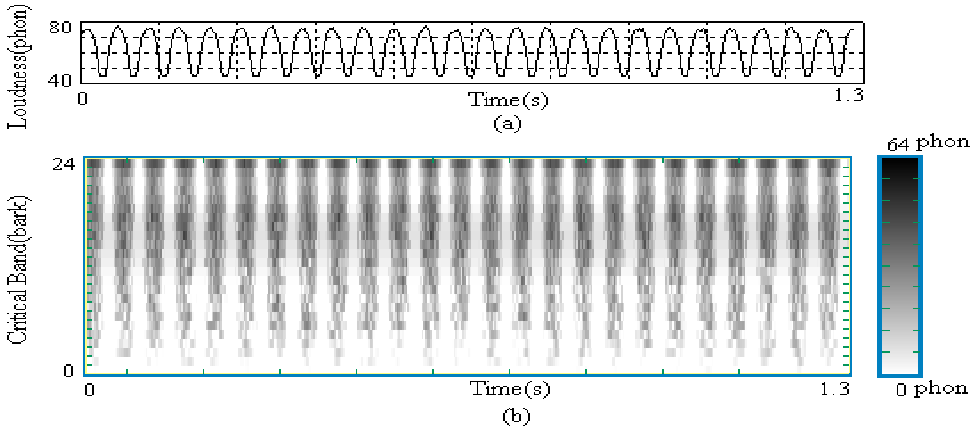

Figure. 10 (a) Time varying loudness and (b) time varying critical band loudness of a 20 Hz modulated white noise signal

The time varying total loudness curves in Figs. 9 and 10 reflect the modulated process of the loudness very clearly. It is a good way to use them to estimate the fluctuation strength and roughness.

3. CONCLUSION

A novel time varying loudness algorithm is proposed. The algorithm uses the FFT based complex critical band filter bank to separate the signal into 24 bark bands and to calculate the short time energy from the envelope output of the filter bank. By use of an equal sound pressure level contour that is obtained from the equal loudness level curves, the short time energy is transformed into the time-varying loudness. This result produces the variation of the loudness with time, and can be used to estimate fluctuation strength and roughness. Some examples presented show the advantages of the new algorithm.

REFERENCES

[1] Zwicker and H. Fastl, Psychoacoustics. Facts and Models, 2nd ed., Springer, Berlin-Heidelberg-New York, 1999.

[2] S.S. Stevens, ‘The Measurement of Loudness, Journal Acoust. Soc. Amer., 27, 1955, pp. 815-829.

[3] J. Tackett, MATLAB Central File Exchange - ISO 226 Equal-Loudness-Level Contour Signal, 2005. http://www.mathworks.com/matlabcentral/fileexchange/

[4] ISO226:2003，Acoustics. Normal equal-loudness-level contours，September 2003

[5] G. Prünster, An empirical study on the sensation of roughness, Conference on Interdisciplinary Musicology – Proceedings, Graz, Austria, April 2004

[6] J. J. Chatterley, Sound Quality Analysis of Sewing Machines, masters thesis, Brigham Young University, August 2005.

[7] J. Timoney, et al, Implementing Loudness Models in MATLAB, Proc. of the 7th Int. Conference on Digital Audio Effects, Naples, October 2004.