Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Auditory Wavelet Transform Based On Auditory Wavelet Families

Y.Salimpour, M.D.Abolhassani , *Member, IEEE*

***Abstract*—The auditory periphery system receives a one dimensional acoustical signal that describes how the local pressure varies with time. However, this one dimensional signal information is then somehow unfolded into a two dimensional time-frequency plane, that tells us when which frequency occurs. The hearing process is based on compromise between time localization and frequency localization. A kind of time-frequency or wavelet type transformation is done in auditory signal processing. In this study the similarities between auditory transform based on the auditory physiological process and wavelet transform are introduced. Specially, band pass filter bank properties and variable time and frequency resolutions with the signal frequency are considered. The main goal is to find the scaling function while the numerical values of the wavelet function were measured. If the wavelet function and the scaling function from the measured data are estimated, then the wavelet coefficients and the scaling coefficients could be calculated. Therefore, the multiresolution implementation of auditory based wavelet transform is possible.**

I. Introduction

IN the inner ear or cochlea, sound is detected by an array of several thousand hair cells that transduce the mechanical vibrations into electrical activities. The cochlea is often thought of as a bank of filters because it performs frequency analysis using a frequency to place mapping along the basilar membrane. That is, each place along the membrane has a characteristic, for which it is maximally displaced when a pure tone of that frequency is presented as an input.

The individual hair cells, and the auditory nerve fibers to which they are connected, are tuned to specific frequencies[1]. The population of the auditory nerve fibers thus provides us with a frequency analysis of sound waveforms in the environment. Each auditory nerve fiber may be considered as a filter that signals information about the temporal structure of the stimuli within their preferred frequency range. As engineers have understood for years, the design of a filter involves an inevitable trade-off between the precision of frequency tuning and temporal tuning. A tone consists of cyclical fluctuations of air pressure, and to obtain an accurate frequency estimate, many cycles must be integrated. However, a longer integration period means a decrease in the temporal accuracy of the filter. In other words, a filter cannot signal both the frequency and the timing of a sound with arbitrary precision. Yet discrimination of the real-world sounds often require accurate measurements of both frequency and timing. Precise temporal information is also important for the sound localization, which in many cases depends on time-of-arrival differences between the two ears. The challenge for the auditory system, then, is to find the right trade-off between timing and frequency analysis [2].

Yousef Salimpour is PhD student in Computational Neuroscience in School of Cognitive Neuroscience. Institute for Studies in Theoretical Physics and Mathematics at Tehran, Iran. He is also research assistant in research centre for science and technology in medicine ,Tehran University of Medical Science ,Tehran, Iran (E-mail: : salimpour@ipm.ir).

M.D.Abolhassani is Associate Professor Medical Physics & Biomedical Engineering Department Tehran University of Medical Sciences and Deputy Head of Research Centre For Science & Technology in Medicine (E-mail: abolhasm@sina.tums.ac.ir)

II. Physiological Observation

Physiological observation of tuning curves of auditory periphery indicates their band pass filter banks behaviour (Fig. 1). In comparison with wavelet filter bank it could be concluded that both systems are decomposing input signal into different frequency bandwidth and the coefficient of band pass filters are considered as a representation of the signal. The frequency response of the tuning curves indicates that, like wavelet mother function and daughter functions each frequency response of tuning curve could be obtained by shifting and translation of the certain tuning curve frequency response[1][3].

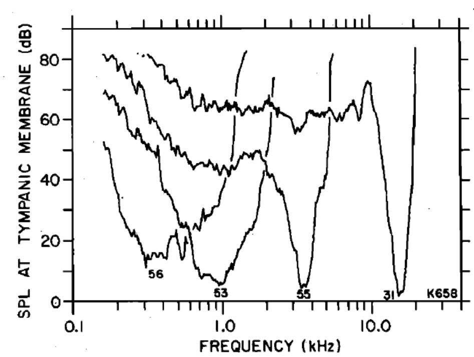


Fig. 1. Sample tuning curves for four single units in the auditory nerve of one cat (from Kiang [1]).

III. The Wavelet transform as an approach to cochlear filtering

It is well known from Fourier theory that, a signal can be expressed as the sum of a possibly infinite series of sines and cosines. This sum is also referred to as a Fourier expansion. The big disadvantage of a Fourier expansion however, is that it has only frequency resolution and no time resolution. This means that although we might be able to determine all the frequencies present in a signal, we do not know when they are present. To overcome this problem in the past decades, several solutions have been developed which are more or less able to represent a signal in the time and frequency domains at the same time. The wavelet transform or wavelet analysis is probably the most recent solution to overcome the shortcomings of the Fourier transform. In wavelet analysis the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and for every position, the spectrum is calculated. Then this process is repeated many times with a slightly shorter (or longer) window for every new cycle. At the end the result will be a collection of time-frequency representations of the signal, all with different resolutions. Because of this collection of representations we can speak of a multiresolution analysis. In the case of wavelets we normally do not speak about time-frequency representations but about time-scale representations. Scale is in a way, the opposite of frequency, because the term frequency is reserved for the Fourier transform. The wavelet analysis described in the introduction is known as the *continuous wavelet transform* or *CWT*. More formally, it is written as:

$$\gamma(s,\tau) = \int f(t)\psi^{*}_{s,\tau}(t)dt \quad (1)$$

This equation shows how a function $f(t)$ is decomposed into a set of basis functions $\psi_{s,\tau}(t)$, called the wavelets. The variables s and τ, scale and translation, are the new dimensions after the wavelet transform. The wavelets are generated from a single basic wavelet $\psi(t)$, the so-called *mother wavelet*, by scaling and translation.

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi(\frac{t-\tau}{s}) \quad (2)$$

In (2) s is the scale factor, τ is the translation factor and the factor $s^{-1/2}$ is for energy normalization across the different scales. Later in this paper, It will be shown that the transform which is done in the cochlea could be estimated by kind of wavelet transform. The two main roles of the cochlea are to separate the input acoustic signal into overlapping frequency bands, and to compress the large acoustic intensity range into the much smaller mechanical and electrical dynamic range of the inner hair cells. When a sound wave hits our eardrum, the oscillations are transmitted to the basilar membrane in the cochlea. The cochlea is rolled up like a spiral; imagine unrolling it (and with it, the basilar membrane), and putting an axis x on to it, so that points on the basilar membrane are labeled by their distances to one end (for simplicity, we use a one dimensional model neglecting any influence of the transverse direction on the membrane or its thickness)[4]. Therefore, the variations of the air pressure at the ear are mechanically transferred in to movement of the basilar membrane which is located in the cochlea. The basilar membrane is equipped with hair cells that react on deviation of the membrane from its rest position. If the cochlea is imagined unrolled the basilar membrane extend along real axis. Sound information at any point can be represented as real function B(x,t), the deviation of the membrane inside the cochlea at position x and time t. Experimental measurements show that for a sinusoidal stimulus the response might be as express in equation (3) which it means the dependency of ψ_ω is approximately a logarithmic shift ,Therefor for input acoustic signal f(t) which has got its Fourier Transform F(ω) can be obtained as follows[11]:

$$B(x,t) = \psi_\omega(x-\log(\omega))e^{j\omega t} \quad (3)$$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)e^{j\omega t}d\omega \quad (4)$$

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-j\omega t}dt \quad (5)$$

By substituting ***f(t)*** in (3) and putting (4) as f(t), ***B(x,t)*** function could be written as follow:

$$B(x,t) = \frac{1}{2\pi}\int F(\omega)\,\psi_\omega(x-\log(\omega))e^{j\omega t}d\omega \quad (6)$$

$$B(x,t) = \frac{1}{2\pi}\iint f(\tau)e^{-j\omega\tau}d\tau \; \psi_\omega(x-\log(\omega))e^{j\omega t}d\omega \quad (7)$$

By some mathematical manipulation of (7) following term will be appear:

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}\psi_\omega(x-\log\omega)e^{j\omega(t-\tau)}d\omega \quad (8)$$

By considering ***x=-log s,*** the final equation for basilar membrane movement could be written as:

$$B(t,-\log s) = \int f(t)\frac{1}{s}\psi'(\frac{t-\tau}{s})dt \quad (9)$$

Therefore, the transform of acoustic signals from the eardrum to the cochlea with a logarithmic scale along the basilar membrane could be approximated by continuous wavelet transform.

$$B(t,-\log s) = |s|^{-1/2}W_\psi(t,-\log s) \quad (10)$$

In the above equation W_ψ is the continuous wavelet transform as defined by (1). In this sense the cochlea can be seen as a natural wavelet transformer[7] .

Physiological observations justify that the auditory system has got sort of wavelet like transform behaviour. Neural tuning is measured by measuring the spiking activity in an auditory nerve fiber as a function of the frequency and intensity of a search tone. The locus of threshold intensities that cause a neuron to fire slightly above its spontaneous rate is called the neural tuning curve. The superscript indicates that the probe intensity is at threshold. Each neuron has such a tuning curve, which is tuned to its "best" characteristic frequency.[8]

IV. IMPLEMENTATION OF AUDITORY WAVELET TRANSFORM

First of all, the whole construction is based on a continuous wavelet transform. In practice this is of course a discrete but very redundant transform, heavily oversampled both in time and in scale. In order to be practical, we need a fast implementation scheme ,This was achieved by nonredundant wavelet bases, by using subband filtering schemes. If a function f(x) is continuous, has null moments, decreases quickly towards zero when x tends towards infinity, or is null outside a segment of range, it is a likely candidate to become a wavelet. It seems that the tuning curves values satisfy the above conditions. In this case we have some physiological measurements or tuning curves

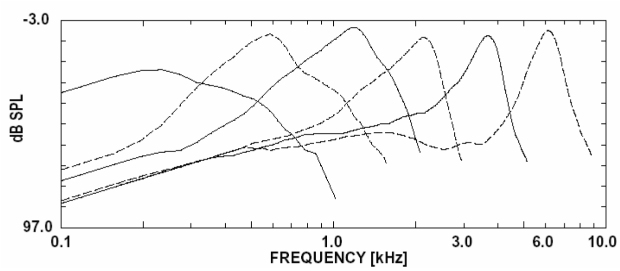


Fig.2. Cat neural tuning curves from Eaton Peabody Lab. The pressure scale, in dB, has been reversed to make the curves look like filter transfer functions. The response "tail" for the 6 kHz neuron is the "flat" region between 0.1 kHz and frequency In the tail the sound must be above 65 dB SPL (which on this scale is down) before the neuron will respond .

(Fig. 2).The numerical value of the curves (Fig. 3) were used as an approximation of the mother wavelet (Fig. 4). By using the definition of scaling function the approximation of scaling function is also calculated (Fig. 5).

Scaling function: when $W\psi\ (t,s)$ is known only for $s<s_0$, to recover original function we need a complement of information corresponding to $W\psi\ (t,s)$ for $s<s_0$, this is obtained by introducing a scaling function $\varphi(t)$ that is an aggregation at scales larger than one[6].

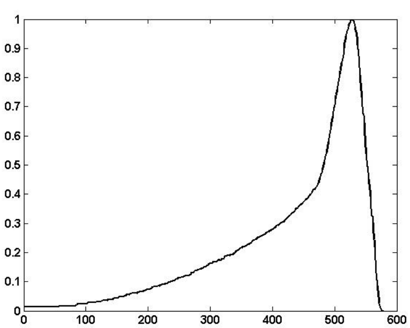


Fig.3..Numerical value of normalized tuning curve which was measured from cat's auditory nurve fibre.

The modulus of its fourier transform is defined by the folowing equation (11),[12]:

$$\left|\phi(\omega)\right|^2 = \int_{\omega}^{+\infty} \frac{\left|\psi(\xi)\right|^2}{\xi} d\xi \qquad (11)$$

So in our representation the problem is to find the scaling function while the numerical values of the wavelet function were measured. If we can estimate the wavelet function and the scaling function from the measured data, the wavelet coefficients and scaling coefficients are calculated (12),(13),(14),(15).

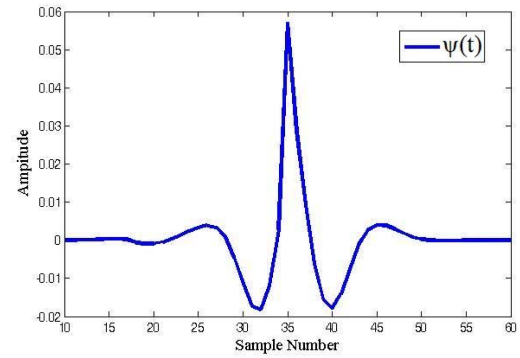


Fig. 4. Wavelet function for auditory wavelet transform

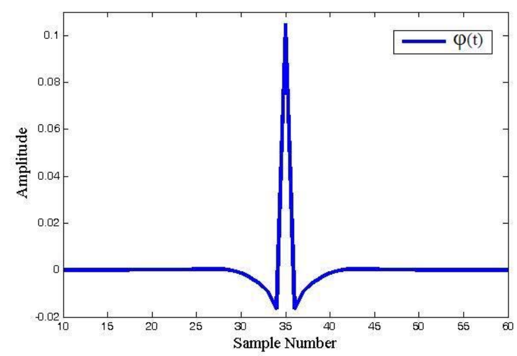


Fig. 5. Scaling function for auditory wavelet transform

$$\Phi(\omega) = \prod_{k=1}^{\infty} \frac{1}{\sqrt{2}} G_0(e^{j\omega/2^k}) \qquad (12)$$

$$\Psi(\omega) = \frac{1}{\sqrt{2}} G_1(e^{j\omega/2}) \prod_{k=2}^{\infty} \frac{1}{\sqrt{2}} G_0(e^{j\omega/2^k}) \qquad (13)$$

$$\Phi(\omega) = \prod_{k=1}^{\infty} \frac{1}{\sqrt{2}} H_0(e^{-j\omega/2^k}) \qquad (14)$$

$$\Psi(\omega) = \frac{1}{\sqrt{2}} H_1(e^{-j\omega/2}) \prod_{k=2}^{\infty} \frac{1}{\sqrt{2}} H_0(e^{-j\omega/2^k}) \qquad (15)$$

If the wavelet coefficients and scaling coefficents were caculated (Fig. 6), the multiresolution impelemtation of auditory based wavelet transform is possible.The analysis stage and the synthesys stage of multiresolution implementation are accomplished (Fig. 7). Now the complete realization of auditory wavelet transform is posible.However first of all we must choose proper wavelet function and by means of that the scaling function ,scaling coefficent and wavelet coeficient can be calculated[4],[6].

V. RESULT

As it was shown the wavelet transform performs the log-linear frequency analysis and the constant quality factor, and can be used as an approximation of auditory acoustic signal transform. The cochlear impulse response was used for choosing the analyzing wavelet transform. The impulse response at 20mm from the oval window was selected as an wavelet function, because its peak frequency is about 1000Hz and in log- linear scale this frequency is almost at the center of the audible range (Fig. 3). The scaling function from a selected mother function is calculated, and by means of these two function the wavelet coefficients and scaling coefficients are calculated (Fig. 6). By this consideration an auditory wavelet transform could be realized (Fig. 7).

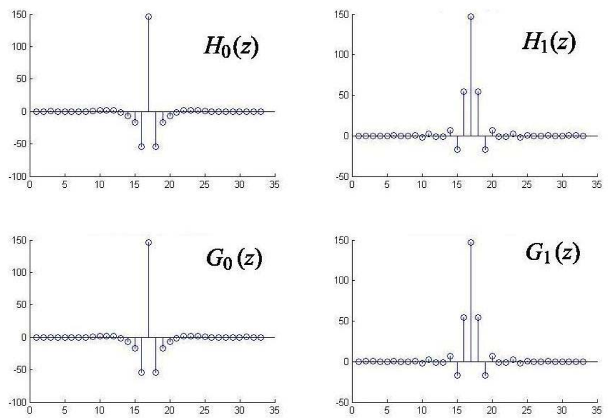
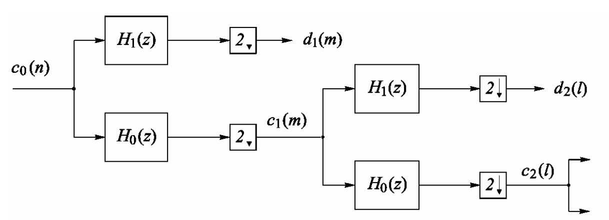
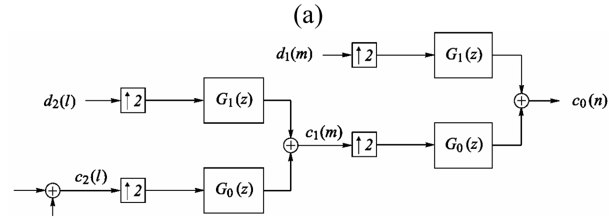


Fig.6. Scaling Coefficients and wavelet coefficients function for auditory wavelet transform



(a)



(b)

Fig. 7. (a) Analysis stage of multiresolution decomposition implementation of auditory wavelet transform .(b)Synthesis stage of multiresolution implementation of auditory wavelet transform.

VI. DISCUSSION

A comparison between the auditory periphery acoustic signal transform and the wavelet transform shows that, although there are similarities specially in band-pass filter bank property and variable time and frequency resolution with the signal frequency, the experimental measurements show that there are some differences and the main difference is in the quality factor. The wavelet transform is a filter bank with constant quality factor, but the physiological research in the hearing system found the quality factor in some frequency band is changing and highly influenced by the activities of the hair cells.

VII. CONCLUSION

The wavelet transform perform the log-linear frequency analysis with constant quality factor filtering and can therefore simulate the auditory model. The cochlear tuning curve is used for choosing the analyzing wavelets or mother function which determines the overall filter shape. The impulse response at medium distance from oval window could be chosen as an analyzing wavelet because its peak frequency is almost at the center of audible range. Since in the realization of the auditory wavelet transform the selection of bases function done by the response of cochlea, therefore it could be considered as a transform in auditory periphery signal processing and must be applicable in audio and speech processing, cochlea implant and specially in the otoacoustic emission signal processing[9],[10].

REFERENCES

[1] N.Y. S. Kiang, ‘’Processing of speech by the auditory nervous system.”*J. Acoust. Soc. Am,* vol. 68, no. 3, pp. 830-835, 1980.

[2] N. Y. S. Kiang and E. C. Moxon, “Tails of tuning curves of auditory nerve fibers,” J. *Acoust.* Soc. *Am,* vol. *55,* no. 3, pp. 620-630, 1974.

[3] P. M. Sellick, R. Patuzzi, and B. M. Johnstone. “Measurement of basilar membrane motion in the guinea pig using the Mossbauer technique.” Journal of the Acoustical Society of America, vol.72, pp. 131-141, 1982.

[4] T. Irino and R. D. Patterson, “A time-domain, level-dependent auditory filter: the gammachirp," *J. Acoust. Soc. Am.*, vol. 101, no. 1, pp. 412{419, January 1997.

[5] T.Irino and R. D. Patterson, “A compressive gammachirp auditory filter for both physiological and psychophysical data," *J. Acoust. Soc. Am.*, vol. 109, no. 5, pp. , May 2001

[6] T. Irino and M. Unoki, “An analysis/synthesis auditory filterbank based on an IIR implementation of the gammachirp," *J. Acoust. Soc. Jap.*, vol. 20, no. 5, pp. 397{406, November 1999.

[7] C. d'Alessandro. Auditory-based wavelet representa-tion. In M. Cooke, S. Beet, and M. Crawford, editors, “Visual representations of speech signals”, pp.131-137. John Wiley & Sons Ltd., 1993.

[8] Y. Salimpour, M.D. Abolhassani and H. Soltanian-Zadeh "AUDITORY WAVELET TRANSFORM" The 3rd European Medical and Biological Engineering Conference, November 20 -25, 2005, Prague, Czech Republic.

[9] Y.Salimpour, Mohammad.D. Abolhassani, A.Ahmadian, K.Barin “Multiresolution Analysis of Transient Evoked Otoacoustic Emission” Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005.

[10] M. D. Abolhasani, Y.Salimpour ,S.Sarkar, "Reproducibility Enhancement of Otoacoustic Emission on Based on Multirate Signal Processing", MEDICON and HEALT TELEMATICS 2004, Xth Mediterranean Conference On Medical and Biological Engineering, August 2004, Italy.

[11] A.Aldroubi,M.Uncer,” Wavelets in Medicine and Biology ”,CRC press ,Inc. pp.528-529 ,1996.

[12] S.Mallat ,“A Wavelet Tour of Signal Processing”, Academic Press,1999.