

# TIME-FREQUENCY MAPPING BASED ON NON-UNIFORM SMOOTHED SPECTRAL REPRESENTATIONS

*P.Hatziantoniou<sup>1</sup>, D.Tsoukalas<sup>1</sup>, J.Mourjopoulos<sup>1</sup>, S.Salamouris<sup>2</sup>*

<sup>1</sup>Audio Group, Wire Communications Lab.,  
Electrical & Computer Eng. Dept.,  
University of Patras, Patras 26500, GREECE

<sup>2</sup>Roister Speakers  
145 Sappous str., Kallithea,  
Athens 17675, GREECE

## ABSTRACT

For audio and acoustic system responses, the auditory system tends to smooth high frequency spectral regions and to register, mainly, low-frequency components of late echoes in the time domain. To model such processing, a theoretical model for non-uniform spectral smoothing is introduced allowing the choice of arbitrary frequency resolution profiles to match such auditory features. This generalized framework is extended to allow mapping of such smoothing spectra into time domain.

## 1. INTRODUCTION

The last decades have seen an impressive research activity in the evolution of alternative time-frequency analysis [1] and multiresolution signal representation methods [2].

However, although the tools established through such work (Fourier and related Transforms, STFT, the Wavelet transform, etc), are suitable for many applications, there still exists a large class of cases for which neither of the existing joint time-frequency nor of the time-scale analysis methods are appropriate and new, flexible processing tools must be developed [3]. The problem of the non-uniform frequency resolution of the auditory mechanism has already attracted significant work for such alternative time-frequency methods, allowing analysis of audio/acoustic signals in warped-frequency scales via FFT and z-Transform [4],[5], time-dependent frequency warping via Wavelets [6], non-uniform filter Banks, and fractional octave Transforms [7]. The principle motivation behind such work is that the auditory mechanism interprets signals at reduced frequency resolution with increasing frequency and due to this for over half century now, engineers represent the frequency response of audio/acoustic systems in log frequency scale, often at a fractional octave resolution (e.g. as 1/3 octave average). Such representation is more recently supplemented by more advanced ERB or Bark scale representations [8] (see Fig. 1).

The work reported here is based on similar motives and addresses the practical problem of appropriate time-frequency modification of audio (e.g. loudspeaker) and acoustic (e.g. room) responses in order to accommodate such auditory features. Especially in such highly dispersive multipath systems as rooms, the ear tends to largely ignore the high frequency components of late reflections, whilst registering more mid and low frequency regions of such late response components. This feature is implemented as a magnitude spectrum modification in the MLSSA measurement system [9] (Fig. 1).

Here, to address such problems, a generalized framework is introduced which allows the choice of any arbitrary and psychoacoustic spectral resolution profile and to map such operations into the time domain via the introduction of a novel variable with frequency window function. All subsequent analysis of the discrete-time audio/acoustic response functions  $h(n)$ , sampled at a frequency  $f_s$  (Hz), will refer to finite duration data, practically implemented via the initial application of a (half) window  $w(n)$  of duration  $N$  (samples), whose length will also determine the lowest response frequency  $f_L$  (Hz) which can be represented in an unambiguous way. Hence, the corresponding complex frequency response  $H(k)$ , will be practically bounded between  $f_L$  (Hz) and the folding frequency  $f_s/2$  (Hz).

## 2. SPECTRAL SMOOTHING

Let us first consider the simplified spectral smoothing operation (weighting), starting from the (windowed) response Transfer Function  $H(k)$  where  $k$  is the discrete frequency index ( $0 \leq k \leq N-1$ ). Then, the complex smoothing operation may be described as a circular convolution:

$$H_{sm}(k) = H(k) \circledast W_{sm}(k) = \sum_{i=0}^{N-1} H((k-i) \bmod N) \cdot W_{sm}(i) \quad (1)$$

where the symbol  $\circledast$  denotes the operation of circular convolution and  $W_{sm}(k)$  is a low-pass filter function. For simplicity here, this function can be supposed to have an ideal low-pass filter form according to the following equation:

$$W_{sm}(k) = \begin{cases} 1/(2m+1) & , \quad k=0, \dots, m \\ 1/(2m+1) & , \quad k=N-m, \dots, N-1 \\ 0 & , \quad \text{otherwise} \end{cases} \quad (2)$$

where  $m$  is the sample index corresponding to the cut-off frequency  $f_c$  (Hz), according to the expression  $m = (N/f_s) \cdot f_c$ . Given the band-limited nature of  $H(k)$ , the above operation does not practically generate any overlapping artifacts on  $H_{sm}(k)$ . In general,  $W_{sm}(k)$  should be arbitrary and complex, but the above expressions represent it as real, assuming it to be a zero-phase function. This assumption was made due to physical considerations, since with smoothing it is required to avoid imposing any unwanted effects on the phase of the original

function. From eq. (1) and given that  $h(n)/H(k)$  and  $w_{sm}(n)/W_{sm}(k)$  constitute Fourier pairs, then:

$$N \cdot h(n) \cdot w_{sm}(n) \xleftrightarrow[N]{DFT} H(k) \circledast W_{sm}(k) \quad (3)$$

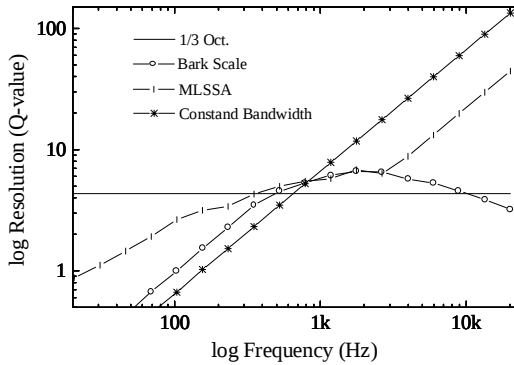
Based on the ideal smoothing function of eq.(1), the corresponding time window function  $w_{sm}(n)$  can be easily evaluated as:

$$w_{sm}(n) = \left( \frac{1}{N(2m+1)} \right) \frac{\sin((\pi n/N) \cdot (2m+1))}{\sin(\pi n/N)} \quad (4)$$

It is now necessary to define a general non-uniform operator allowing the required variable with frequency smoothing. Given that the "quality" factor  $Q=f/\Delta f$  [10], traditionally describes the smoothing properties of filters, then the discrete-frequency version of this function will be  $Q(k)=k \cdot f_{bin}/\Delta f_k$ , where  $f_{bin}$  gives the DFT bin separation and  $\Delta f_k$  indicates the filter bandwidth for any value of the frequency index  $k$ . Then the discrete variable  $m$  may be expressed as a function of  $k$  by the following equation:

$$m(k) = \{(k \cdot f_{bin})/Q(k)\} \cdot c \quad (5)$$

where  $c$  is a normalization factor, hence allowing a variable degree of spectral averaging for each value of the discrete frequency,  $k$ . To allow  $W_{sm}(k)$  to accommodate the variation of  $m(k)$  implied by eq. (5) and given it must also depend on  $k$  (eq.(2)), it is necessary to express it as the more general function  $W_{sm}(m(k),k)$ , for  $m(k)=1, \dots, (N/2)-1$ . Such an expression allows flexible adaptation into the traditional fractional octave smoothing, when  $W_{sm}(m(k),k)$  has constant value over a range of values of  $k$ , and bandwidth increasing with frequency (Fig. 1).



**Figure 1.** Resolution vs. frequency of various existing frequency analysis schemes.

The function  $W_{sm}(m(k),k)$ , representing all the possible cut-off combinations of  $W_{sm}(k)$  as are implied from the previous expressions, can be then presented as a  $M \times N$  matrix  ${}^2W_{sm}$ , where  $M < N/2$  is the maximum value for the function  $m(k)$ . Each row of this matrix represents a frequency vector for the smoothing

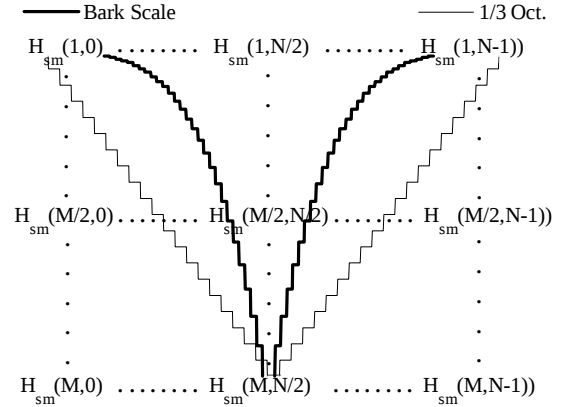
filter, for a specific value of  $m(k)$  and each column represents all the possible cut-off values of the filter at each discrete frequency  $k$ . From eq.(1), the general form of the desired non-uniform smoothing will be now given as:

$$H_{sm}(m(k),k) = W_{sm}(m(k),k) \circledast H(k) \Leftrightarrow \Leftrightarrow H_{sm}(m(k),k) = \sum_{i=0}^{N-1} W_{sm}(m(k),i) \cdot H((k-i) \bmod N) \quad (6)$$

In order to evaluate eq.(6),  $H(k)$  must be defined for all values of  $i$ , and this allows it to be represented by a  $N \times N$  complex matrix  ${}^2H$ , where each row is derived by the circular shifting of the  $H(k)$  elements (modulo  $N$ ). So, the matrix form of non-uniform spectral complex smoothing will be:

$${}^2H_{sm} = {}^2W_{sm} \cdot {}^2H = \begin{bmatrix} \sum_{i=0}^{N-1} W_{sm}(1,i) \cdot H(0-i) & \dots & \sum_{i=0}^{N-1} W_{sm}(1,i) \cdot H(N-1-i) \\ \vdots & \ddots & \vdots \\ \sum_{i=0}^{N-1} W_{sm}(M,i) \cdot H(0-i) & \dots & \sum_{i=0}^{N-1} W_{sm}(M,i) \cdot H(N-1-i) \end{bmatrix} \quad (7)$$

By choosing (tracing)  $N$  elements following specific paths in the 2-D space described by  ${}^2H_{sm}$ , it will be possible to derive the desired function  $H_{sm}(k)$ , which will be a non-uniform spectrally smoothed function, derived from  $H(k)$  (see Fig. 2).



**Figure 2.** Derivation of smoothed spectral vector by tracing through the  ${}^2H_{sm}$  matrix.

More formally such an operation will be described as follows. Let  $U_k$  be a  $N \times N$  matrix with elements  $u_{ij}$  such that  $u_{ij}=1$ , for  $i=j=k$  and  $u_{ij}=0$  for all other values of  $i$  and  $j$ . Now, let  $v_m$  be a  $1 \times M$  vector, having its  $m$ -th element equal to 1 and all other equal to 0. Then,  $v_m \cdot {}^2H_{sm} \cdot U_k$  will generate a  $1 \times N$  vector with all elements equal to 0, except of the  $k$ -th which will be equal to  $H_{sm}(m(k),k)$ . Given that  $k$  has values in the range  $[0, N-1]$ , then it may possible to generate  $N$  different vectors in the previously described way, which when summed will produce the required 1-D smoothed sequence. Since  $m$  may be derived from a general

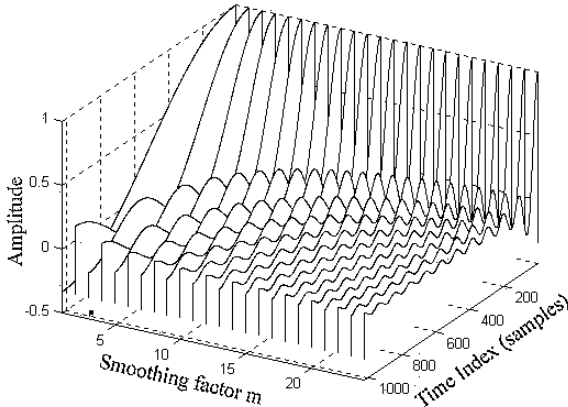
frequency/resolution function  $m(k)$  (see eq.(5)), then the required vector of the smoothed sequence, i.e.  ${}^1\mathbf{H}_{sm}$  will be obtained as:

$${}^1\mathbf{H}_{sm} = \sum_{k=0}^{N-1} (\mathbf{v}_{m(k)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_k) \quad (8)$$

### 3. FREQUENCY TO TIME MAPPING

It is also possible to approach the non-uniform spectral smoothing processing via time domain, by considering variable windowing of  $h(n)$ . Let us at first consider the window function  $w_{sm}(n)$ , which it may have the typical form defined by eq. (4). Then in order to accommodate the different cut-off  $m(k) \in [1, N/2)$  of the corresponding spectral smoothing function, this window may be represented by the function  $w_{sm}(m(k), n)$

which it can be presented in a matrix form as  ${}^2\mathbf{w}_{sm}$ , where each row represents a time vector for the window for a specific value of  $m$  and each column represents all the possible variations of the window at each discrete value of the time index  $n$  (Fig. 3). It can also be easily deduced that the rows of matrices  ${}^2\mathbf{w}_{sm}$  and  ${}^2\mathbf{W}_{sm}$  constitute Fourier Transform pairs.



**Figure 3.** Window functions corresponding to different smoothing factor

As was the case with spectral smoothing, it is also possible to define  $h_{sm}(m(k), n)$  as a sequence which describes the 2-dimensional windowing of the signal sequence  $h(n)$ . That is:

$$h_{sm}(m(k), n) = N \cdot h(n) \cdot w_{sm}(m(k), n) \quad (9)$$

Note that, by definition,  $h_{sm}(m(k), n)$  &  $H_{sm}(m(k), k)$  are Fourier pairs, only for each specific value of the function  $m(k)$ . It is now necessary to map the spectral smoothing operation into the time domain. In the previous section equation (8) described the non-uniform spectral complex smoothing of  $H(k)$  as the summation of vector elements in the  $M \times N$  matrix  ${}^2\mathbf{H}_{sm}$ . Given also that  $m(k) = m(N-k)$ , then equation (8) can be also expressed as:

$${}^1\mathbf{H}_{sm} = (\mathbf{v}_{m(0)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_0) + (\mathbf{v}_{m(N/2)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_{N/2}) + \sum_{k=1}^{(N/2)-1} [(\mathbf{v}_{m(k)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_k) + (\mathbf{v}_{m(k)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_{N-k})] \quad (10)$$

The summation described in eq. (9) contains  $(N/2)+1$  vector terms, each of  $N$  elements. Let  $\mathbf{S}_k$ ,  $k=0, \dots, N/2$ , be each one of these vectors, i.e.

$$\mathbf{S}_k = \begin{cases} \mathbf{v}_{m(k)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_k + \mathbf{v}_{m(k)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_{N-k} & k \neq 0, k \neq N/2 \\ \mathbf{v}_{m(0)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_0 & k = 0 \\ \mathbf{v}_{m(N/2)} \cdot {}^2\mathbf{H}_{sm} \cdot \mathbf{U}_{N/2} & k = N/2 \end{cases} \quad (11)$$

For convenience, it is possible to define the factor  $E_N$  as the complex number  $E_N = e^{j2\pi/N}$ . Thus the IDFT  $N \times N$  matrix  $\mathbf{E}$  for an  $N$ -point IDFT is built from the powers of  $E_N$ . Then the IDFT of vector  ${}^1\mathbf{H}_{sm}$  will yield the required time-domain sequence of the smoothed spectrum, i.e.  ${}^1\mathbf{h}_{sm}$ , that is:

$${}^1\mathbf{h}_{sm} = \frac{1}{N} {}^1\mathbf{H}_{sm} \cdot \mathbf{E} = \frac{1}{N} \left( \sum_{k=0}^{N/2} \mathbf{S}_k \right) \mathbf{E} = \sum_{k=0}^{N/2} \left[ \frac{1}{N} (\mathbf{S}_k \cdot \mathbf{E}) \right] \quad (12)$$

The vector term  $(1/N) \cdot (\mathbf{S}_k \cdot \mathbf{E})$  inside the sum is the IDFT of the vector  $\mathbf{S}_k$  yields a time vector  $\mathbf{s}_k$ . It is easy to show that the  $n$ -th element of the vector  $\mathbf{s}_k$ ,  $s_k(n)$  will be given by:

$$s_k(n) = \frac{2}{N} [h_{sm}(m(k), n) \odot \cos(2\pi kn/N)] \quad (13)$$

Then, from eqs. (12) and (13) the  $n$ -th element of the vector  ${}^1\mathbf{h}_{sm}$ ,  $h_{sm}(n)$  will be:

$$h_{sm}(n) = \frac{2}{N} \sum_{k=0}^{N-1} (h_{sm}(m(k), n) \odot \cos(2\pi kn/N)) \quad (14)$$

Finally, from eq.(9) and eq.(14), the non-uniform spectral smoothing processing via time domain is defined as:

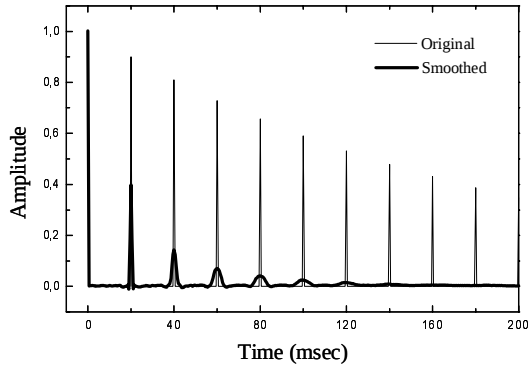
$$h_{sm}(n) = 2 \cdot \sum_{k=0}^{N/2} (w_{sm}(m(k), n) \cdot h(n) \odot \cos(2\pi kn/N)) \quad (15)$$

or equivalently:

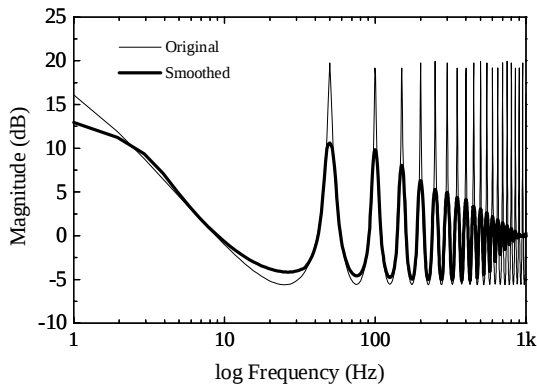
$$h_{sm}(n) = 2 \cdot \sum_{k=0}^{N/2} \sum_{\ell=0}^{N-1} \left[ (w_{sm}(m(k), \ell) \cdot h(\ell)) \cos\left(\frac{2\pi k}{N}(n-\ell)\right) \right] \quad (16)$$

### 4. RESULTS

The above process was successfully applied to many practical audio/acoustic response measurements, but the results can be more clearly illustrated for a characteristic dispersive system response, such as a comb-filter. Typical results of the non-uniform smoothing in both time and frequency domains of a comb-filter are shown in Fig. 4. As can be observed, late reflection components are progressively suppressed and increasingly low-passed without any time-scale modification, and similarly, the corresponding spectrum is progressively smoothed resulting to a reduction of the characteristic comb-like profile.



(a)

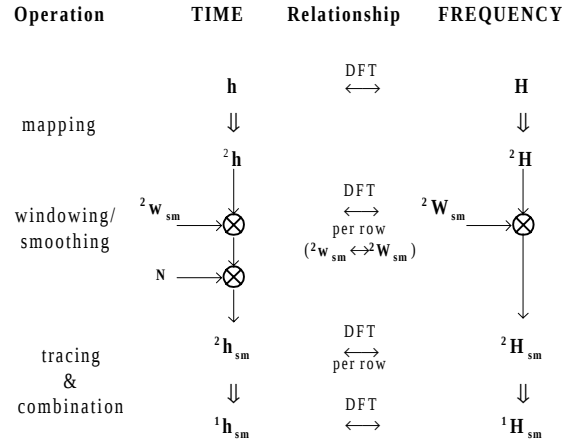


(b)

**Figure 4.** Original comb-filter sequence and its non-uniform spectrally smoothed version; (a) time domain sequences, (b) magnitude spectra.

## 5. CONCLUSIONS

Finite-length and band-limited data sequences such as audio/acoustic system responses can be processed by arbitrary or psychoacoustically-derived frequency smoothing profiles so that smoothed versions may be derived both in frequency and time domains, as was explained. The procedure for such operations via either frequency or time domain, it is summed in Fig. 5, generally conforming to the framework described in [3]. From this figure it can be deduced that windowing/smoothing is implemented as matrix operations between the mapped data matrix and the required window/filter matrix, from which the smoothed data matrix is derived. To obtain the required smoothed sequence, an appropriate tracing and combination procedure must be applied on the 2-D smoothed data. This method introduces a very flexible, efficient and generalised processing procedure, appropriate for many audio/acoustic applications. Finally, by observing eq. (16), it can be noted that the form of the smoothed time-domain operation, bears some relation to that of MDCT [11], but significantly, there are also the following differences: (a) each time sample is calculated using the summation over  $k$  of sums having the general form of MDCT, and (b) the window function within the "MDCT-like" sum depends on the function  $m(k)$ .



**Figure 5.** Time-frequency mapping procedure for the smoothed spectral representation.

## 6. ACKNOWLEDGEMENT

This work was funded by the program "YPER-113" of the Greek General Secretariat for Research and Technology.

## 7. REFERENCES

- [1] Cohen L. "Time-Frequency Distributions - A Review", *Proc. of the IEEE*, 77.7, (941-981), 1989.
- [2] Vetterli M. and Herley C., "Wavelets and Filter Banks: Theory and Design", *IEEE Trans. on Signal. Processing*, 40.9, (2207-2232), 1992.
- [3] Baraniuk G.R. and Jones D.L. "Unitary Equivalence: A New Twist on Signal Processing". *IEEE Trans. on Signal Processing*, 43.10, (2269-2282), 1995.
- [4] Oppenheim A. V., Johnson D. H., and Steiglitz K., "Computation of Spectra with Unequal Resolution Using the Fast Fourier Transform", *Proc. IEEE*, 59, (299-301), 1971.
- [5] Karjalainen M., Harma A., Laine U.K., "Realizable Warped IIR Filters and Their Properties", *Proc. IEEE ICASSP-97*, (Munich) Apr. 1997.
- [6] Evangelista G. and Cavaliere S., "The Discrete Time-Frequency Warped Wavelet Transforms", *Proc. IEEE ICASSP-97* (Munich) Apr. 1997.
- [7] Martin N., Mars J., Martin J., and Chorié C., "A Capon's Time Octave Representation Application in Room Acoustics", *IEEE Trans. on Signal Processing*, 43.8, (1842-1854), 1995.
- [8] Karjalainen M., Piirilä E., Järvinen A. and Huopaniemi J. "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques", *Preprints of 102nd AES Convention*, (Munich), 1997, Preprint 4437.
- [9] Rife D. D., "MLSSA (Maximum-Length Sequence System Analyzer), Reference Manual, v.9", *DRA Laboratories*, 1994.
- [10] Brown J.C., "Calculation of a Constant Q Spectral Transform", *J. Acoust. Soc. Am.*, 89.1, (425-434), 1991.
- [11] Princen J., Johnson A., and Bradley A., "Subband/Transform Coding Technique Based on Time Domain Aliasing Cancellation", *Proc. IEEE ICASSP*, IEEE Press, Piscataway, N.J., 1987, pp.2161-2164.