NON-ORTHOGONAL GABOR REPRESENTATION OF BIOLOGICAL SIGNALS

Mark L. Brown[1] *William J. Williams*[2] *Alfred O. Hero III*[2]

[1]Bioengineering Program
[2]Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, MI 48109, USA
e-mail: mlbrown@eecs.umich.edu

ABSTRACT

A new technique for modelling biological signals as a linear combination of non-orthogonal Gabor logons is described. The technique has been applied to two types of signals, Event-Related Potentials (ERPs) and temporomandibular joint (TMJ) clicks. Examination of time-frequency representations of these signals revealed that they appear to consist of a small number of localized energy concentrations. Attempts to capture this apparent low dimensionality with the standard orthogonal Gabor expansion and the standard wavelet transform were unsuccessful. However, the non-orthogonal Gabor decomposition method described in this paper provides a compact, accurate signal representation and the parameters provide a good basis for ERP category and TMJ click classification.

1. INTRODUCTION

Signal processing techniques based on a stationary or time-invariant signal model are frequently inadequate for the analysis of biological signals. However, such techniques are often applied to Event-Related Potentials (ERPs). Signal averaging methods discard potentially useful individual record variation under the inaccurate assumption that the ERP is composed of a stationary deterministic signal in uncorrelated additive noise. Fourier analysis, often applied in EEG analysis, cannot capture in a meaningful way the transient nature of an ERP.

Time-frequency analysis, particularly the Exponential Distribution (ED), has been successfully employed by Choi, Williams and Zaveri[1] to gain an understanding of ERPs. An improved and generalized version of the ED, the Reduced Interference Distribution (RID)[2] continues to be used in these studies. Features of the time-frequency representation of these signals have been identified which distinguish ERPs according to stimulus category. They were able to discriminate categories on the basis of energy values at five points in the time-frequency representation of an ERP. Details of the experimental methods and significance of the categories will be given in the next section.

While the techniques employed by Choi, Williams and Zaveri provide good signal classification, they do not provide a good signal representation since there are an infinite number of signals with the required energy values at the five locations in time-frequency space. Thus, the features of the ERP which distinguish category provide only a glimpse of the over-all time and frequency structure of the ERP. What we now seek is a compact signal representation which retains the information necessary to reconstruct the entire signal and which contains the category specific

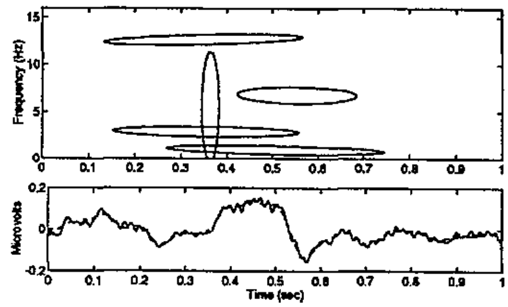

Figure 1. Lower plot shows an ERP (solid) line and the model fit (dashed line). The correlation coefficient between the model and data is .98. The upper plot shows the time and frequency location, chirp rate and spread of the 5 logons composing the model. The contours are located at 1/2 the peak value of the logon.

information necessary for pattern classification. It is expected that greater insight into the underlying dynamics of the generating system will be achieved with this new representation.

In observing time-frequency representations of the ERPs, we noticed that the ERPs appeared to be composed of a small number of high energy features localized in time and frequency. We decided that an ERP might be efficiently represented as a linear combination of time and frequency shifted Gabor logons. A parametric technique such as this has the advantage over the non-parametric time-frequency methods of decreasing noise sensitivity at the expense of increased bias. The non-orthogonal Gabor representation method that has resulted is accomplishing the goals of modelling the data well and characterizing the ERP categories. Figure 1 illustrates how well the data are modelled.

Temporomandibular joint (TMJ) sounds were recorded during examination of subjects with suspected pathological conditions. The sounds recorded are indicative of the type and severity of the pathology. Recent work on the analysis of these sounds by Widmalm, Williams and Zheng [3] has involved the use of time-frequency distributions (TFDs). The TFDs reveal a time-frequency energy distribution which is compact in time and frequency, the type of signal efficiently represented by Gabor logons. Type 1, 2 and 3 joint clicks are easily modelled with a linear combination of one, two or three logons. Figure 2 shows a type 1 joint click with the model fit and its TFD.

0-7803-1775-0/94 $3.00 © 1994 IEEE

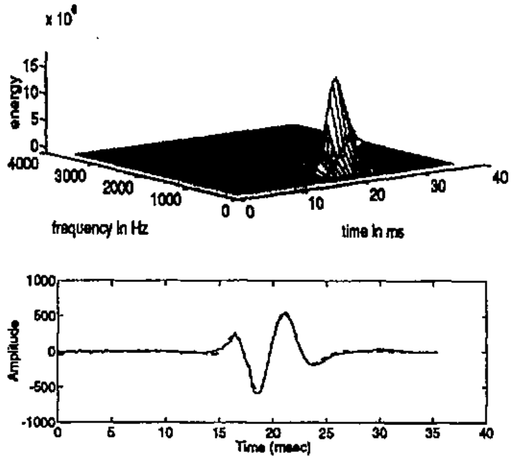

Figure 2. Lower plot shows a type 1 click TMJ sound (solid line) and the model fit using a single logon (dashed line). The correlation coefficient between the model and data is .99. The upper plot shows the TFD of this joint click.

2. EXPERIMENTAL METHODS

The ERP data for this study were collected from clinical patients at the Riverview Psychiatric Out-patient Clinic[4]. Patients in the study suffered from phobias or pathological grief reactions. Words presented at a 3-40 ms exposure via tachistoscope were selected from 4 categories: Category 1 words were chosen from words used by the patient to describe his/her symptom; Category 2 words were chosen by clinicians to be related to the emotional causes of the complaint; Category 3 words were selected from Osgood's list of pleasant words and Category 4 words were selected from Osgood's list of unpleasant words. 32 words, 8 from each category, were presented in random order 6 times each to the patient. Data from six electrodes were digitized at 250 Hz and filtered from 0.1 to 70 Hz. Only data recorded from a midline electrode one third of the distance from Cz to Pz have been used to date for this study. Each ERP consists of 256 data samples, 100 ms pre-stimulus and 900 ms post-stimulus.

The TMJ sounds were recorded with a heart sound microphone fastened to the subject's forehead while the subject opened and closed his mouth. The recordings were performed at the University of Michigan School of Dentistry. The sounds were sampled at 7200 Hz.

3. NON-PARAMETRIC SIGNAL REPRESENTATION

Numerous techniques exist for representing well-behaved signals in different domains. Many of these techniques, such as wavelets, principal components, Karhunen-Loeve, etc., can give insight into the signal which is often not obvious when viewing the signal in the time-amplitude domain. Most of these techniques involve representing the signal as a weighted linear combination of vectors from an appropriately chosen orthonormal basis. The weights then form the representation of the signal in the new domain. The basis is usually chosen to highlight features of the signal and perhaps to separate a desirable part of the signal from an undesirable part of the signal (such as to separate signal from noise). If successful, this results in groups of large-valued weights separated by small-valued weights. That is, the resulting representation localizes signal features. The most familiar representation or transformation is the Fourier transform which provides a frequency representation of the signal. When the spectral content of a signal stays relatively constant and when different components of the signal are well separated by spectral composition, the Fourier transform is a powerful tool of analysis. When the spectral content of the signal varies appreciably in the observation window, the Fourier transform fails to capture this dynamic. Time-frequency and time-scale representations have been developed to provide a meaningful picture of the spectral dynamics of a time-varying signal. The basis functions for these representations are localized in time-frequency or time-scale space, respectively. A number of desirable properties for these representations have been identified, among them time support, frequency support and preservation of the marginals. Also desirable is completeness of the basis so that all signals in the vector space of admissable signals, usually L_2 or l_2, can be fully represented as a linear combination of these vectors.

One such representation is the standard orthogonal Gabor transform. Gabor[5] showed that any signal in L_2 could be represented as the weighted sum of modulated and shifted Gaussian functions (logons) centered on a rectangular lattice in time and frequency under the constraint that $T\Omega \leq 2\pi$ where T is the time sampling interval and Ω is the frequency sampling interval. That is,

$$s(t) = \sum_{m,n=-\infty}^{\infty} C_{m,n} h_{m,n}(t) \tag{1}$$

where

$$C_{m,n} = \int_{-\infty}^{\infty} s(t)\gamma^*_{m,n}(t)dt \tag{2}$$

$$h(t) = \frac{1}{\sqrt[4]{\pi\sigma^2}} e^{-t^2/2\sigma^2} \tag{3}$$

$$h_{m,n}(t) = h(t - mT)e^{jn\Omega t} \tag{4}$$

$$\gamma_{m,n}(t) = \gamma(t - mT)e^{jn\Omega t} \tag{5}$$

$\gamma(t)$ is derived from $h(t)$ such that

$$\frac{2\pi}{\Omega T} \int_{-\infty}^{\infty} h(t)\gamma^*_{m,n}(t)dt = \delta(m)\delta(n). \tag{6}$$

The standard orthogonal Gabor transform appeared to be a good candidate for representing ERPs because of the compact structure of the logons in time-frequency space and because of the similarity in appearance of the logons to components of the ERP signal. Figure 3 shows the resulting representation for an ERP. We desire a signal representation for ERPs that needs only a small number of terms to describe the signal so that further analysis of category-specific signal features is simplified. As figure 3 demonstrates, this representation does not meet our requirement for describing the signal with a small number of terms. Apparently, this basis does not adequately describe this signal for our purpose.

Friedlander and Porat[6] provide some insight into what is wrong with the standard orthogonal Gabor transform

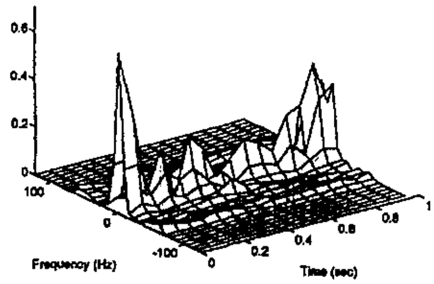

Figure 3. Standard orthogonal Gabor transform of an ERP showing the magnitudes of $C_{m,n}$.

method for representing ERPs. They have developed an algorithm which uses the orthogonal Gabor expansion for detecting transient signals. They showed that when the analyzing window, a damped sinusoid in their work, matches the signal well, and the time-frequency lattice points match the time of occurrence and central frequency of the signal, the standard orthogonal Gabor representation does an excellent job of localizing the signal at these lattice points. However, when the shape of the analysis window is not parsimonious with the signal or the signal events do not occur at discrete points on the time-frequency grid, many more terms may be required to adequately represent the signal with this technique. Looking at figure 1 we see that the shape of the logon which best fits the ERP data is different at different points in the time-frequency plane. In consideration of the work of Friedlander and Porat, it seems reasonable to account for the large number of significant terms in the standard orthogonal Gabor transform, observed in figure 2, as a result of a mismatch between the signal and the analysis window in addition to a mismatch between the position of the analysis grid points and the signal.

The standard wavelet transform provides a similar time-frequency description to that of the Gabor transform with an important difference in the way in which the time-frequency plane is tiled or covered by the basis functions. Equations (1) and (2) which describe the Gabor transform are equally desriptive of the standard wavelet transform. However, the subscripts m, n desribe not the time and frequency shifts, but the time and scale shifts of the analysis window.

$$h_{m,n}(t) = a_0^{-m/2} h\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \quad (7)$$

Figure 4 provides a schematic description of the tiling of the time-frequency plane by the standard wavelet transform when $a_0 = 2$, which is a common implementation. As scale of the wavelet increases, the frequency increases and the duration decreases. Comparing the shape of the Gabor logons depicted in figure 1 to the tiling of the time-frequency plane by the standard wavelet transform as depicted in figure 4, we see that the pattern of scaling in ERPs does not match well with the standard wavelet transform. For example, the logon centered at t=.36 sec, f=5 Hz in figure 1 has a shorter duration and wider bandwidth than the logon centered at t=.35 sec, f=12.5 Hz. A signal composed of scaled and shifted components which is efficiently

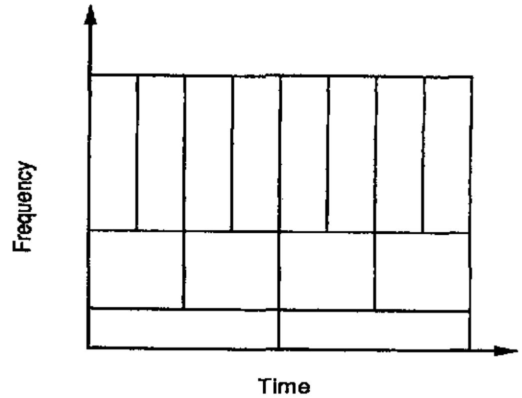

Figure 4. Schematic of the tiling of the time-frequency plane by the standard wavelet transform.

represented by the standard wavelet transform would have just the opposite trend.

4. NON-ORTHOGONAL GABOR DECOMPOSITION

In the orthogonal Gabor transform, orthogonality is achieved by restricting scale and translation parameters to the integers, i.e. to a discrete rectangular lattice. This restriction prevents a parsimonious representation of signals whose major components lie between lattice points. As a particularly extreme example consider the case of figure 2 where the signal can be represented by a single logon with fractional translation and scale parameters while an infinite number of logons would be required in the orthogonal Gabor transform.

We sought a parsimonious representation for the ERPs that was compact in time and frequency and required only a small number of terms to represent the ERP. The best non-parametric techniques described in the previous section did not meet the requirement of a compact representation. Thus, we began to investigate parametric time-frequency representations. The Gabor logon [5] defined in equation (3) was used as the basis signal for our representation. As previously stated, any signal in L_2 can be completely represented as a sum of time- and frequency-shifted logons. Therefore, this representation will possess many of the desirable properties for time-frequency distributions including satisfying the time and frequency marginals. Our goal is to use a smaller set of logons each parameterized in such a way as to match the signal characteristics at different parts of time-frequency space. Thus, we reduce the dimension of the space needed to represent the signal, but with the assurance that we can completely represent any signal with enough logons. Our expectation was that a small number of non-orthogonal logons would adequately represent ERPs and TMJ sounds. Consequently, we began to explore a method of fitting the data with a linear combination of 4 or 5 Gabor non-orthogonal logons each parameterized by 6 values: central time, central frequency, spread, amplitude, chirp rate and phase. Figure 5 shows the TFD of an ERP which displays a very distinctive Gabor logon shape. Figure 6 shows the time domain and time-frequency description of a Gabor logon with increasing frequency.

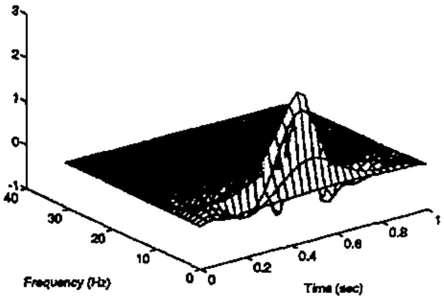

Figure 5. TFD of an ERP which has distinctive Gabor logon shape.

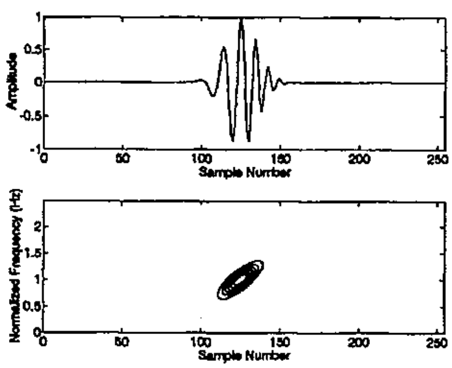

Figure 6. Upper plot shows a Gabor logon with increasing frequency in the time-domain. The lower plot shows the time-frequency description of the same Gabor logon.

4.1. Implementation

The following algorithm was developed to demonstrate the effectiveness of the Gabor decomposition. The algorithm is implemented as follows:

1. Normalize the signal(ERP or TMJ sound) to zero mean, unit energy.
2. For each logon (up to 5 total) exhaustively search the parameter space for the five parameters: center time, center frequency, phase, spread, chirp rate. Amplitude is then found using linear least squares. The parameters are chosen to minimize the sum of square residuals. Thus, the first logon fits to the largest energy concentration in the signal, the second to the next largest concentration and so on.
3. As each logon is identified, it is added to the model signal which is initially zeroed. The model signal is used in the computation of the sum of square residuals.
4. The resulting parameters from the above steps are then used as the starting point of a gradient search fit in an attempt to find a global minimum.

5. RESULTS AND DISCUSSION

ERP data sets from 10 subjects have been fit using the non-orthogonal Gabor representation. The first data fits did not use chirp rate as a parameter, nor was the gradient search phase of the algorithm implemented. However, for 8 of the 10 data sets, the mean correlation coefficient between the data and the model was in the range .85–.90. Examination of the data sets which were not well fit by the model revealed a more complicated time-frequency representation for these subjects. In order to fit the data of these subjects the additional parameter for the logons, chirp rate, is required. Preliminary work with this model indicates that the additional parameter will allow fits for these subjects which are comparable to the other 8.

In addition to evaluating the quality of the fit using the correlation coefficient, the ability to discriminate categories on the basis of the parameters was evaluated. We attempted to measure the ability to discriminate stimulus category by forming a linear discriminant function using odd numbered records of a data set and applying the discriminant function to the remaining data records. This resulted in a classification success as good as that achieved by Choi, Williams and Zaveri indicating that indeed the Gabor decomposition has captured the category characteristics. Classification success was significantly better than chance ($p < .05$) for 4 of the 8 data sets that were well-fit by the fitting algorithm.

The TMJ work is quite recent. Good data fits using a small number of non-orthogonal logons has been achieved as seen in figure 2 and it appears nearly certain that highly reliable automated TMJ sound classification will be possible at least in part on the basis of the logon parameter space.

6. ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation, NSF grant no. BCS 9110571.

7. BIBLIOGRAPHY

[1] Choi, H. I., W. J. Williams, and H. Zaveri, "Analysis of Event-Related Potentials: Time-Frequency Energy Distribution," *Biomedical Sciences Instrumentation*, vol. 23, 1987, pp. 251-258.

[2] Williams, W. J., and J. Jeong, "Reduced Interference Time-Frequency Distributions," Chapter 3 in **Time-Frequency Signal Analysis: Methods and Applications**, B. Boashash (ed.), Longman and Cheshire, Melbourne, Australia, 1992.

[3] Widmalm, S. E., W. J. Williams, C. Zheng, "Time frequency distributions of TMJ sound," *Journal of Oral Rehabilitation*, vol. 18, 1991, pp. 403-412.

[4] Shevrin, H., W. J. Williams, R. E. Marshall, R. K. Hertel, J. A. Bond, L. A. Brakel, "Event-Related Potential Indicators of the Dynamic Unconscious," *Consciousness and Cognition*,no. 1, 1992, pp. 340-366.

[5] Gabor, D., "Theory of Communication," *J. of IEE*, vol. 93, part III, no. 26, 1946, pp. 429-457.

[6] Friedlander, B. and B. Porat, "Detection of Transient Signals by The Gabor Representation," *IEEE Trans. ASSP*, vol. 37, no. 2, 1989.