Analysis of Spontaneous Activity in Cultured Brain Tissue Using the Discrete Wavelet Transform

J.D. Johnson[1*], D. Plenz[2], J. Beggs[2], W. Li[1], M. Mieier[1], N. Miltner[1], and K. Owen[1]

[1]*Bioengineering Department*
The University of Toledo, Toledo Ohio
**jeffjohns@eng.utoledo.edu*

[2]*National Institute of Mental Health*
National Institutes of Health
Bethesda, Maryland

Abstract

Multi-microelectrode array devices can be used to simultaneously record activity from multiple neurons distributed in a tissue slice. One of the brain functions being investigated with microelectrode arrays is the periodic behavior of spontaneously active neurons in the cortex and basal ganglia.. However, these recording methods generate several hundred megabytes of data per hour and, currently, there is no efficient and accurate approach for the identification of the repeated pattern. We present an approach that uses the discrete wavelet transform to accelerate identification of repeating patterns of neural activity. We perform match filtering on the coefficient data, not the time-domain data. Our wavelet approach operates on ¼ the data but provides similar classification abilities as the time domain correlation.

1. Introduction

In animals, learning is a function of repetition. The animal experiences and responds to a cue multiple times before its behavior in response to the cue is sufficiently strengthened. We are interested in understanding this process at the neuronal level. We know that the cue is encoded in the brain as a pattern of cortical activity and that this activity is the input to the striatum, one of the components of the basal ganglia. Within the basal ganglia, learning is hypothesized to involve changes in corticostriatal (cortex-to-striatum) synapses and is a function of the cortical activity and subsequent dopaminergic activity from the substantia nigra (another component of the basal ganglia). In other words, cortical activity provides the input to the basal ganglia and the neurotransmitter dopamine is the training signal that shapes the response of the striatum to that activity.

From microelectrode array technology, Plenz and Kital [1] propose that subthalamic nucleus (STN) and globus pallidus external (GPe) form a synchronized oscillation. The absence of dopamine gives rise to the commonly observed tremors experienced by patients with Parkinson's disease. Abeles et al [2,3] and Shao and Tsau [4] have developed several statistical methods to show spatial and temporal correlation in the firing patterns of simultaneously recorded neurons. However, they simplified the recorded data to binary data. That is, the data is treated as one to indicate the neuron is firing and zero to indicate the neuron is not firing at sampling point. It is apparent that much information is lost and artifact noise is added into the data by these approaches.

1.1 Organotypic Tissue Cultures

Unfortunately, it is extremely difficult to record from the necessary sites in the basal ganglia in an awake and behaving animal. The task of precisely placing dozens of electrodes in the brain, each within 50 μms of their targeted site, is a technological challenge that has not yet been overcome. Thus, Dr. Dietmar Plenz, Chief of the Unit of Neural Network Physiology, National Institute of Mental Health (NIMH), at the NIH, is using a procedure to create tissue cultures that mimic the organization of the basal ganglia. They place tissues from the cortex, striatum, and substantia nigra on a glass slide and incubate for four weeks. During incubation, the tissues extend processes (dendrites and axons) into the neighboring tissues as they would during normal brain development. Thus, after culturing, the tissue displays the same overall physiologic organization as the basal ganglia displays. While limited behavioral correlations can be made between the intact basal ganglia and these tissue slices, the fact that we have complete access to these cultures provides the opportunity to collect data during the learning process. To record from this tissue, we will employ a microelectrode array technology that consists of a grid of 60 electrodes spaced 200μm apart (Figure 1). The array is placed underneath the cortical component of the culture and we are thus able to simultaneously record from 60 sites across the cortex.

Proceedings of the Third IEEE Symposium on BioInformatics and BioEngineering (BIBE'03)
0-7695-1907-5/03 $17.00 © 2003 IEEE

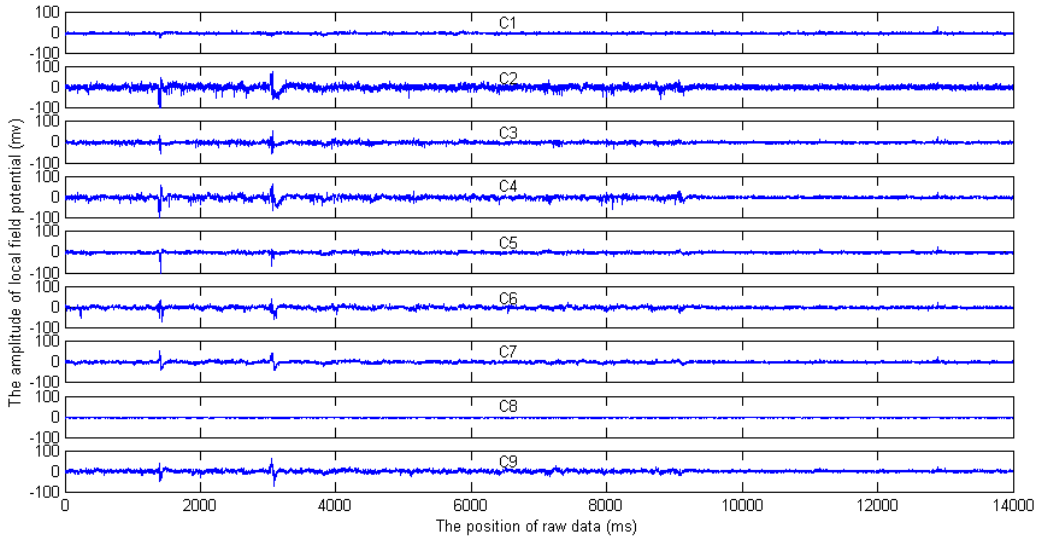

Figure 1: Subset of 60-channel multielectrode neural activity data. This figure shows 9 channels of local field potential deflections in the first 14000 ms of the data file. Sampling is at 1KHz. For this study, we selected nine channels that represented a wide range of signal data found the complete set of sixty channels.

However, in the tissue culture, cortical activity does not encode the presence of a learning cue as it would in vivo because there are no sensors (eyes, ears, etc.) associated with the culture. Fortunately, the cortical component of the culture is constitutively active, that is, it exhibits spontaneous and repeated neural activity. We plan to interpret patterns of spontaneous activity in the cultured cortical tissue as if those patterns encoded the presence of learning cues. Furthermore, if a specific pattern of cortical activity repeats itself, we will interpret this reoccurrence as signaling that the cue corresponding to that pattern also repeated itself. Thus, if we can identify the set of patterns that a tissue spontaneously generates and then recognize the reoccurrence of those patterns, the opportunity arises that allows for the training of the tissue culture through repetition. In the proposed paradigm, training the tissue consists of either directly applying dopamine to the striatal component of the tissue or exciting the dopminergeric neurons in the substantia nigra.

1.2 Real-time pattern classifier system

Before the organotypic tissue cultures can be used to study the neuronal basis of learning, the problems of classifying the spontaneously generated patterns and recognizing the reoccurrence of those patterns must be solved. This is a significant problem because of the amount of data generated by the tissue: approximately 450 MB (million bytes) of data per hour and over 5 GB (billion bytes) of data during a culture's 12 hour lifetime. Complicating the problem is the fact that each tissue produces a unique set of activity patterns so that the patterns found in one tissue will not help in the process of finding patterns in a second tissue. A second complication is the relatively short viable lifetimes of the cultures. The problem of classifying and recognizing patterns cannot occur until recording from the tissue begins. However, once recording begins, only 12 hours elapses before the tissue begins to degrade. The classification and recognition problems must be solved quickly enough to still allow time to conduct extensive experiments on the tissue before it loses viability. Finally, for these cultures to be used in learning experiments, there must be a very short period of time between recognition of a repeated pattern and excitation of the tissue. For biological reasons, the interval between the pattern being generated in the cortical tissue and the dopamine arriving at the striatial tissue must be less than 10 ms. Our system must work within this time constant.

Johnson et al have developed a signal processing approach that is designed to process, in real-time, large amounts of signal data [7]. This approach forms the basis for developing a real-time pattern classifier of spontaneous cortical activity in organotypic tissue cultures. A pattern is defined as a sequence, in time, of spatially located firing activities as recorded by an array of electrodes residing underneath the tissue (See Figure 2). Our approach will be to use Wavelet decomposition to

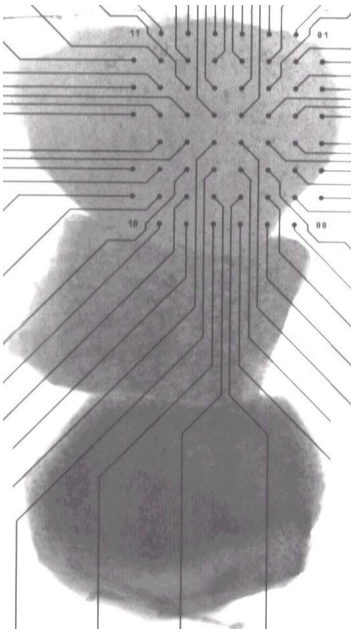

Figure 2. Photograph of tissue slices that make up in vitro basal ganglia. Image shows cortical (top), striatal (middle), and nigra (bottom) tissues after 2 days in vitro. System cultured for 4 weeks to allow dendritic and axonic processes to extend from each tissue into its neighbor. Notice the 60-element microelectrode array situated underneath the cortical tissue.

accelerate the process of identifying repeating patterns of neural activity.

1.3 Discrete wavelet transform

The wavelet transform [5] is a method for complete time-frequency localization for signal analysis and characterization. In comparison to Fourier transform (FT), discrete wavelet transform (DWT) cannot only provide frequency information, but also the time information. The wavelet transform of a signal is its decomposition on a family of real orthonormal bases $\psi_{mn}(x)$ obtained through translation and dilation of a kernel function $\psi(x)$ known as the mother wavelet

$$\psi_{mn}(x) = 2^{-m/2}\psi(2^{-m}x - n) \qquad (1)$$

where $m, n \in Z$, are a set of integers.

Using the orthonormal property of the basis functions, wavelet expansion coefficients of a signal $f(x)$ can be computed as

$$c_{mn} = \int_{-\infty}^{\infty} \psi_{m,n}(x) f(x) dx . \qquad (2)$$

The signal can be reconstructed from the coefficients as

$$f(x) = \sum_m \sum_n c_{m,n} \psi_{m,n}(x) . \qquad (3)$$

In general a mother wavelet can be constructed using a scaling function $\phi(x)$ which satisfies the scaling equation

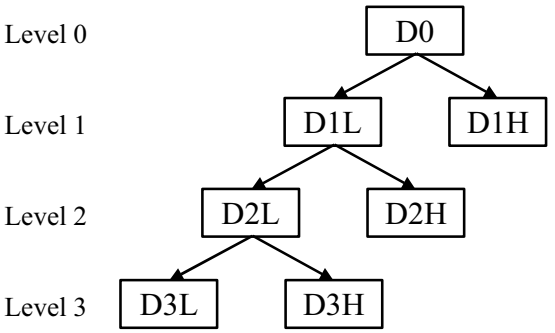

Figure 3: Three level multilevel decomposition tree using the wavelet transform. The original signal is Level 0 (D0). Low frequency (L) and high frequency (H) components are generated after each transform. In this application, only the low frequency component at each level, D*x*L, is further decomposed.

$$\phi(x) = \sum_k h(k)\phi(2x - k) \qquad (4)$$

and the corresponding wavelet defining equation

$$\psi(x) = \sum_k g(k)\phi(2x - k) \qquad (5)$$

where $g(k) = (-1)^{1-k} h(1-k)$. The coefficients of the scaling equation $h(k)$ must satisfy several conditions for the set of basis functions to be unique, orthonormal, and have a certain degree of regularity. For the filtering operations, $h(k)$ and $g(k)$ coefficients can be used as the impulse responses corresponding to the low and high pass operations. As an $O(N)$ algorithm, the discrete wavelet transform of a given sampled signal is a fast algorithm for computing the wavelet expansion coefficients of the signal. The DWT does not work over all possible scales and locations but by choosing the scales and locations of the wavelet based on the powers of 2, the accuracy of the algorithm can still be maintained while the computational time is significantly reduced [6].

Johnson et al [7] has used the DWT to identify an airframe's flutter boundaries during flight-testing, in real-time. The wavelet transform decomposes the response

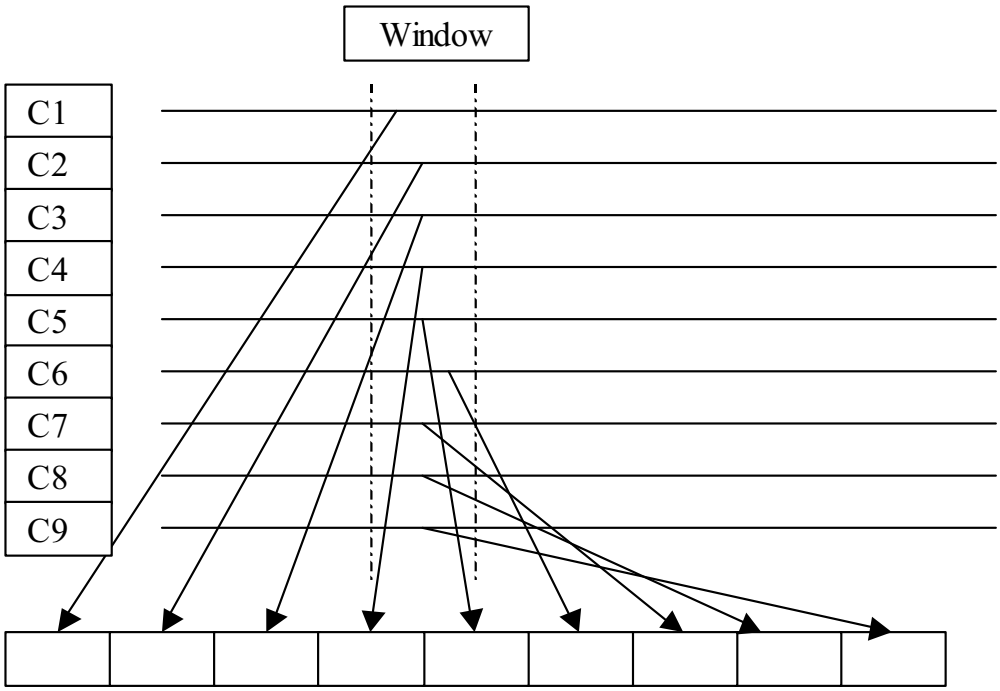

Figure 4: Creation of the data signal. This figure shows how to create a *data signal*. C1 through C9 stands for the individual electrode data channels, Channel 1 through Channel 9. Data from the same 80ms time interval in each channel is concatenated to form a single dimensional data signal.

signal into a set of sub-signals that correspond to different frequency bands. The same operation is applied to each entry in a dictionary of singlet functions. The transform results in a considerable reduction in the data and thus to a reduction in the computational time needed to calculate the correlation. We first create a dictionary of time-domain singlet functions, which include large amount of possible wave functions similar to the airframe's flutter boundaries. The discrete wavelet transform is applied to each entry in the dictionary to obtain a wavelet-domain sub-dictionary. Multiple levels of decompositions produce a new wavelet-domain dictionary in which the size of each of its entries is significantly reduced compared to the entries in the original dictionary. Similarly the discrete wavelet transform is applied to the time-domain response signal to obtain the wavelet-domain sub-signal.

The original dictionary can be viewed as the Level-0 dictionary (D0) in the multi-resolution decomposition tree of the discrete wavelet transform as shown in Figure 3. We apply the wavelet transform to the Level-0 dictionary and create the sub-dictionaries D1L and D1H. The L and H symbolize the low and high frequency components of the transformed data, respectively. For this study, we only utilize the low frequency component of each transform because the high frequency components lacked discriminatory information for our specific task. Sub-dictionary D2L and D2H are obtained by performing the discrete wavelet transform on D1L. The discrete wavelet transform can be further applied to the low-frequency component of each sub-dictionary entry to further decompose the dictionary and reduce the number of the points that represent each entry. In this way, we create a multi-resolution decomposition tree that includes the sub-dictionaries D1L, D2L, and D3L. All the sub-dictionaries can be obtained off-line once the parameter space has been discretized.

By choosing different scales for the wavelets, the wavelet transform can achieve any desired resolution in time or frequency. Each application of the DWT increases the frequency resolution of the signal. Thus, the sub-components at node *N* of the DWT decomposition tree have a finer frequency resolution than the sub-components at the node *N*-1, one level above *N*. Finer frequency resolution implies better frequency localization allowing us to accurately locate the most prominent frequencies in the original signal. The ability of wavelet

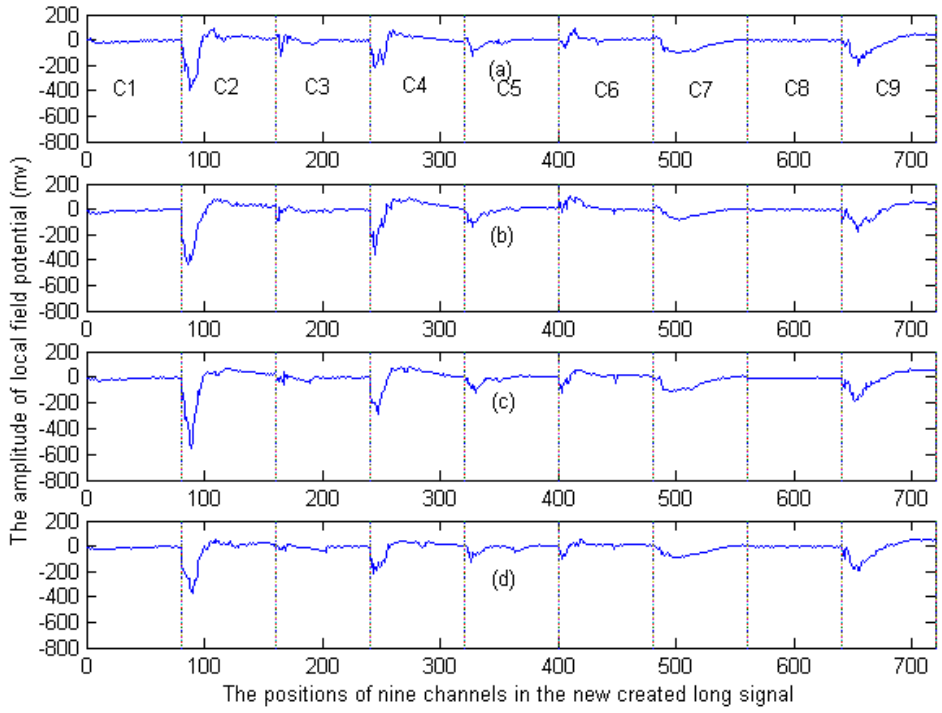

Figure 5: This figure shows a time-domain pattern found to occur four times in the sample file of 400000 ms of data. The regions labeled C1 through C9, indicate that the data in these regions comes from Channels 1 through 9, respectively. The channel data is concatenated to form a single record as shown here. Each channel contributes 80 ms of data starting at the same point in time for all channels. The pattern in (a) occurs at *t* =75,266 ms of the original data, (b) at *t* =163,238 ms, (c) at *t* = 273,714 ms, and (d) at *t* = 325,412 ms. Thus, the first and fourth occurrences of the pattern were separated by approximately 4 minutes (250 seconds).

transforms to operate locally, that is extract frequency characteristics in window of a time-based signal, is the significant difference between wavelet transforms and Fourier transforms

The discrete wavelet transform is also applied to the sampled signal received from the vehicle under test. A multi-resolution decomposition tree is generated to obtain the sub-signal in each level of resolution. The method is the same as that applied to the entries in the dictionary and its sub-dictionary described above.

Time-based correlation is computationally demanding. Thus, instead of computing the correlation between the sampled signal and each entry in the original dictionary in the time-domain, our method computes the correlation coefficients between the sub-signal and each entry of the sub-dictionary in the lowest level of the decomposition tree (Figure 3).

2. Matched Filtering in the Wavelet Domain

The criterion for grouping in the time domain approach consists of the correlation between the coefficients of the incoming signal with existing library entries. If the correlation is greater than the threshold value of 0.9, then the incoming signal will be assigned into the entry that has a correlation coefficient greater than 0.9. Otherwise, a new entry will be created in the library. The same method was used in wavelet approach except that we use a threshold to 0.93. The reason for raising the threshold in wavelet approach is because we are using the low-low frequency sub-band of the wavelet decomposition. Thus, the lack of detail included in this signal requires a more stringent correlation test.

The data signal is created by the following approach:

1. We use a window with size of 80 ms to hold the data from nine channels.

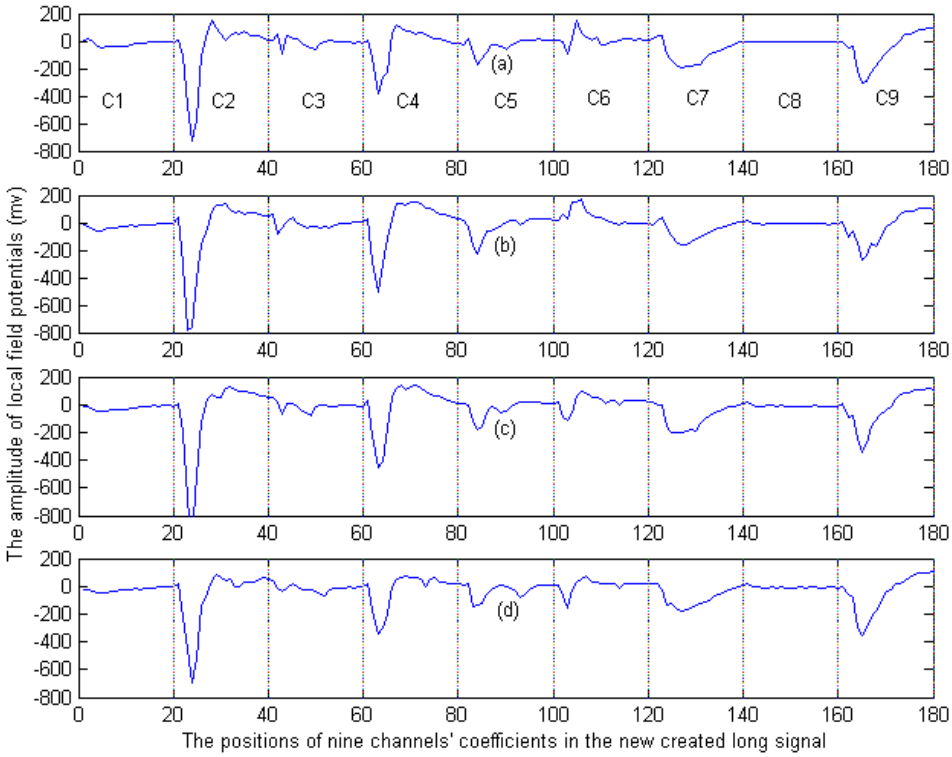

Figure 5: This figure shows a discrete wavelet transform (DWT)-domain pattern found to occur four times in the sample file of 400000 ms of data. The DWT-domain pattern is generated from the wavelet coefficients by doing a two level decomposition on the signal shown in figure 1 and by choosing the low frequency sub-band of the 2nd level decomposition. So the number of coefficients from each channel is reduced to ¼ of the length of original data. The regions labeled C1 through C9 indicate that the data in these regions comes from Channels 1 through 9, respectively. The channel data is concatenated to form a single record as shown here. Each channel contributes 20 coefficients. The pattern in (a) occurs at *t* =75,266 ms of the original data, (b) at *t* =163,238 ms, (c) at *t* = 273,714 ms, and (d) at *t* = 325,412 ms. Thus, the first and fourth occurrences of the pattern were separated by approximately 4 minutes (250 seconds).

2. We concatenate the nine channels into a single vector of dimension 720 (=80×9) elements.
3. This new vector is called a data signal, which is correlated, in the time-domain, against the library entries (identified patterns).

This process can be shown in the Figure 4.

For the wavelet domain, we follow the similar steps as above except we first perform a two-level decomposition before creating the data signal. Also, the data signal is composed of the wavelet coefficients. The length of each entry in the library is 180 (=720 ÷ 4) elements because of the roughly ½ reduction of data in *each* of the levels of the wavelet decomposition.

3. Results

The comparison of sorting results from time domain and 2nd level wavelet decomposition domain

1. Both the time domain and wavelet domain approaches find 71 patterns in a 400 second signal segment.
2. Each approach finds the same set of 71 patterns as evidenced by their identical temporal locations.
3. For the time domain, there are 62 unique patterns and 9 reoccurrences. For the wavelet domain, there are 61 unique patterns and 10 reoccurrences.
4. For the time domain, the 9 reoccurrences are spread across 6 unique patterns. For the wavelet domain, the 10 reoccurrences are spread across 7 unique patterns. Nine reoccurrences in the time domain are spread across the same unique

patterns as nine of the reoccurrences in the wavelet domain.

Figures 5 and 6 show the time- and wavelet-domain results of finding four occurrences of the same pattern. The significant difference in these two approaches is the amount of data each approach is required to operates on. In Figure 5, the time-domain signal is composed of 720 elements. In Figure 6, the wavelet-domain signal is composed of 180 elements. Thus, over the course of match filtering across the entire library of patterns, a significant amount of time is saved in the wavelet domain approach simply because of the significant reduction of data that this approach is required to operate on.

4. Future Work

We presented a preliminary study of the application of the discrete wavelet transform to the problem of finding repeated patterns of activity in cultured brain tissue. We selected a subset of the entire data set based on its representative signal and showed that the wavelet approach has equivalent pattern matching capability compared to the more data intensive temporal approach.

However, for this solution to be realized in real-time, we realize that a significantly greater speed-up of processing must occur. To that end, we have designed our algorithm to run on a Beowulf cluster computer. A Beowulf cluster is a parallel computer composed of a dedicated set of PCs running the multiprocessor Linux operating system. There are several areas where a parallel processing approach can be applied to this problem. As can be seen in Figure 3, the signal decomposition can be implemented in a by organizing the cluster in a binary tree architecture. Each node of the tree will perform a one-level decomposition and pass the result to its two children.

We envision that our final solution will operate in two stages. The first stage will consist of identifying the set of patterns generated by the tissue. Because each tissue generates a unique set of patterns, the first 1-2 hours of data generated by the tissue will be used to identify the patterns specific to the tissue. During this time, we could also apply statistical methods to reduce the number of channels processed by eliminating those channels found to have low information content. The application of parallel computing will be most useful during the first stage of the solution. The product of the first stage of processing would be the wavelet coefficients that characterize each pattern in the set generated by the tissue. The second stage of the solution will be the real-time (<10ms) identification of the reoccurrence of patterns during the remaining viable lifetime of the tissue culture. Our matched filtering approach, using wavelet decomposition is ideally suited for an electronic, chip-level implementation, thus providing identifications well within our required time constraints.

5. References

[1] Plenz, D. and S.T. Kital, *A basal ganglia pacemaker formed by the subthalamic nucleus and external globus pallidus.* Nature, 1999. **400**: p. 667-682.

[2] Abeles, M., *et al.*, *Spatiotemporal firing patterns in the frontal cortex of behaving monkeys.* Journal of Neurophysiology, 1993. **70**(No. 4): p. 1629-1638.

[3] Abeles, M. and G.L. Gerstein, *Detecting spatiotemporal firing patterns among simultaneously recorded single neurons.* Journal of Neurophysiology, 1988. **60**: p. 909-924.

[4] Shao, X.M. and Y. Tsau, *Measure and statistical test for cross-correlation between paired neuronal spike trains with small sample size.* Journal of Neuroscience Methods, 1996. **70**: p. 141-152.

[5] Vetterli, M. and J. Konacevic, *Wavelets and Subband Coding*. 1995, Englewood Cliffs, N.J.: Prentice Hall. 488.

[6] Goswami, J.C. and A.K. Chan, *Fundamentals of Wavelets: Theory, Algorithms, and Applications*. 1999, New York: Wiley.

[7] Johnson, J.D., J. Lu, A. P. Dhawan, R. Lind, *Real-time identification of flutter boundaries using the discrete wavelet transform.* Journal of Guidance, Control, and Dynamics, 2002. **25**(2): p. 334-339.