ABSTRACT

A growing body of evidence from neuroscience and cognitive science suggests that the brain maintains a scale-invariant timeline in which events are represented on a logarithmically-spaced Weber-Fechner scale. This Weber-Fechner representation is scale-invariant – meaning that the array gives the same relative error for signals at any scale – and resource conserving, with the amount of resources necessary to represent a particular timescale going up logarithmically with that timescale. In addition to this time-domain representation, it is in many cases desirable to have a frequency-domain representation with these properties. We show that starting from a logarithmically-spaced timeline, a scale-invariant band-pass decomposition with log-spaced central frequencies can be readily constructed. The method is neurally plausible and can be computed on-line. This approach can be used in neuromorphic engineering to extract spatio-temporal features of interests, for instance slowly varying input components.

Index Terms— Band-pass decomposition, Scale-invariance, Log-spacing, Time-frequency representation

1. INTRODUCTION

Numerous cognitive experiments, but also our everyday experience, demonstrate that humans can remember the time at which events were experienced with gradually decreasing accuracy over a wide temporal range, up to tens or even hundreds of seconds [1, 2, 3, 4]. The gradual decay without abrupt changes in the accuracy suggests that the same mechanism might be used to store the memory over such wide temporal range. It is widely believed that working memory is represented through the maintenance of action potential firing (see for example [5]).

Natural signals, including auditory, visual and natural language signals show correlations over a wide range of time scales [6, 7, 8, 9]. Perhaps as a response to this rich structure, different regions of the human brain contain temporal receptive fields that respond to features across a range of distinct scales [10, 11]. These considerations have led to the hypothesis that at each moment the brain maintains working memory as an estimate of the time at which events took place along a Weber-Fechner scale; this representation can then be used to support a variety of behaviors including representation of time, space and sequences as well as episodic memory [12, 13]. The defining property of the Weber-Fechner representation is that the error in the time dimension grows linearly with the time in the past, consistent with a logarithmically-spaced scale. Intuitively, the difference between 10 and 11 is the same as the difference between 100 and 110. A scale-invariant representation appears to be an adaptive response to a world that contains information at many scales.

Let us consider the problem of spectral decomposition of a continuous time-varying signal \( f(t) \) from the perspective of an organism facing a natural environment. Assume that, like a brain, we do not have access to the entire signal and that the resources necessary to maintain memory for the signal at any moment are costly. If we are to estimate the instantaneous power of the signal in a particular frequency band \( \omega_o \) using standard methods, we would need to maintain on the order of \( \omega_o^{-1} \) time points of the signal. That is, the resources necessary to estimate a particular frequency band grow linearly with the period of that frequency band. In addition, for a continuously-varying natural signal we do not a priori know the frequency (or frequencies) that will be important for our survival. A neurally-plausible time-local method for scale-invariant frequency-decomposition that required logarithmic resources would enable an organism to represent a very wide range of frequencies with minimal resources.

The solution we propose here builds on a neurally-plausible method for constructing a Weber-Fechner timeline [14]. In this method, a set of leaky integrators indexed by their rate constant \( s \) maintain at each moment the Laplace transform of the input signal leading up to that moment. This operation is time-local meaning that the only storage of the signal is through the set of \( s \) values that are maintained. The intrinsic properties of neurons can give rise to a broad range of time scales [15, 16]; these intrinsic properties can be exploited to give rise to leaky integrators with a range of rate constants in a biophysically realistic simulation [17].

The biologically plausible computational method for spectral decomposition presented here operates on the outputs of the neural leaky integrators (effectively low-pass filters that constitute working memory) and constructs a bank of band-
pass filters. The decomposition is scale-invariant and when
the ratio of the adjacent rate constants is a constant it pro-
vides log-spacing of the central-frequencies of the band-pass
filters. The method for constructing such representation is de-
tailed in the next section. We also reflect on the construction
of a Weber-Fechner timeline and show that this band-pass
decomposition yields a Weber-Fechner frequency-line with
analogous properties to the timeline, only in the frequency
domain. Additionally we show that the frequency-line can
also be constructed from the timeline. In the third section
we demonstrate the bandpass decomposition in the presence
of multiplicative noise and contrast its performance to slow
feature analysis (SFA) [18]. In the last section we discuss the
overall properties and potential applications of this method.
We reflect on its possible utility in the context of neuromor-
phic engineering.

2. DERIVING FREQUENCY DECOMPOSITION
FROM NEURAL REPRESENTATION OF THE
RECENT PAST

We will first describe a neural mechanism for maintaining a
scale-invariant working memory and then we will show how
this mechanism can serve as a basis for constructing a neural
representation of time and frequency. The block diagram illus-
trating the basic idea is show on Figure 1. The mechanism
for representing the time is analogous to the one described
in [19], see also [14, 20, 12]. The main contribution of this
paper is in constructing a scale-invariant neural frequency de-
composition with log-spaced central frequencies which can
be used to identify spatiotemporal signal of interest.

2.1. A neural mechanism for maintaining working mem-
ory

We use a set of leaky integrators with different time con-
stants to maintain information about history of a continuously
changing input across multiple seconds. The crucial property
of this representation is that temporal resolution gradually de-
cays from more recent to more distant past.

Time varying input signal $f(t)$ is sent to $N$ nodes constit-
tuting the input layer. Each node is a leaky integrator with a
different time constant $1/s$. The activity of these nodes $F_s(t)$
is described by the following differential equation:

$$\frac{dF_s(t)}{dt} = -sF_s(t) + f(t). \quad (1)$$

The above expression effectively constitutes a set of first or-
der low-pass filters with the cutoff frequencies $\frac{s}{\pi}$ and gain
1/s. This can be read from the transfer function in the Fourier
domain with variable $\omega$:

$$F_s(\omega) = \frac{1}{s} \frac{1}{1 + \frac{s}{\pi}}.$$ \quad (2)

If the number of leaky integrators would be infinite, contain-
ing all the time constants from zero to infinity, the output of
this set would constitute a Laplace transform (restricted to
purely real, positive values of the Laplace variable) of the in-
put with respect to $s$: $F_s(t) = \int_{-\infty}^{t} f(t') e^{-s(t'-t)} dt'$. This
would be a perfect memory representation of the input signal.
Since in practice the number of leaky integrators is limited
one can obtain a discretized approximation of the Laplace
transform by choosing a finite number of neurons (restricting
$s$ to a discrete set of values). Here we chose $\frac{s}{s+1} = C$
and $C > 1$ (exponentially decaying $s$). Log spacing of $s$ is opti-
mal for long range correlated signals ([14], see also [13]). We
will see that it also results in a scale-invariant frequency line.
2.2. Creating a neural timeline

A simple linear transformation on the instantaneous values of the outputs from the leaky integrators is sufficient to construct a neural timeline – a neural memory representation of what happened when. This effectively means that each different stimulus will have a set of neurons that fire sequentially for a circumscribed period of time following each presentation of that stimulus. This representation has two important properties: 1) It is scale-invariant (the standard deviation of the circumscribed period of elevated firing rate is proportional to its peak time); 2) Peak firing times of the sequentially activated cells are log-spaced. These two properties imply that the resolution of the estimate of when a stimulus happened will decay with time elapsed since the stimulus presentation. This is consistent with the memory decay observed in everyday life and quantified in behavioral experiments. Additionally, this is a resource saving property: we need log of the resources that we would need for uniformly spaced peak firing (assuming the same maximal time resolution).

To construct a neural timeline, since \( F_s(t) \) is an approximation of the Laplace transform of the input signal, we need to invert it and express \( F_s(t) \) in terms of some new variable that will correspond to internal representation of time, \( \hat{t} \). This is achieved by setting the weights between the first and the second layer such that the output of the second layer \( \tilde{f}_1(t) \) is approximately a \( k^{th} \) order derivative of \( F_s(t) \) with respect to \( s \) (see [19] for details):

\[
\tilde{f}_1(t) = \frac{(-1)^k}{k!} s^{k+1} F_s^{(k)}(t).
\] (3)

where \( s = -k/\hat{t} \). We define this linear operator as: \( L_k^{-1} \) so that \( \tilde{f}_1 = L_k^{-1}[F_s] \). This inversion will provide a dynamically updated timeline maintained by a population of neurons, such that the firing rate of a particular neuron in the second layer represents the input activity \( \hat{t} \) time ago (Figure 2a and 2b). Neurons with these properties have been reported in multiple areas of the mammalian brain (these are called time-cells, see [21] for a recent review).

To explore the frequency properties of this representation we again compute the transfer function using a Fourier transform. Notice first that the impulse response of \( F_s(t) \) for \( f(t) = \delta(0) \) is: \( F_s(t) = e^{-st} \) and in that case the \( k^{th} \) order derivative of \( F_s(t) \) reads as \( F_s^{(k)}(t) = (-t)^k e^{-st} \). We can now construct a transfer function in the Fourier space by inserting the above expression into equation (3):

\[
\tilde{f}_1(\omega) = \frac{1}{(1 + \frac{\omega}{s})^{k+1}}. \] (4)

The above transfer function is equivalent to a low-pass filter of order \( k+1 \) and unit gain. This property will be useful later to demonstrate that we can construct a frequency-line from this representation as well.

![Fig. 2. Scale-invariant log-spaced neural support across time and frequency dimensions (timeline and frequency-line). The y-axis in all 4 plots represents the firing rate. Both representations are constructed using only instantaneous firing rate from a set of exponentially decaying cells that received an input at time 0. a (time in lin-scale) & b (time in log-scale): Time-cells fire sequentially for a circumscribed period of time following the input at time 0. Each curve represents a single cell with scaled amplitude. The firing fields (intervals with elevated firing rate) spread in a scale-invariant fashion. Peak firing times are log-spaced. This set of cells maintains memory of the previous values of the stimulus such that what happened in a more distant past is represented less accurately than what happened in a more recent past. c (frequency in lin-scale) & d (frequency in log-scale): Similarly to time-cells in the time domain, frequency responses of neurons that do the frequency decomposition span the frequency axis in a scale-invariant fashion with log-spaced central frequencies: slower components are represented more precisely than faster components. Each curve again represents a single cell, but in this case what we plot is a frequency response of the cells to a delta-pulse. This set of cells decomposes the input into ordered frequency bands allowing an estimate of the speed of the input signal.](image-url)
tifying slow components) to estimate the speed of the input signal. This can be easily done by taking a time derivative of the signal. However this approach is prone to errors when the scale of the target signal is not a priori known, as is often the case. Therefore, we will use our dynamical memory representation to construct derivatives on multiple (log-spaced) scales. It turns out that when \( \frac{s_{j+1}}{s_j} = C \) subtracting the activity of neighboring cells in \( F_s(\omega) \), scaled by \( s \) (to have unit gain, Figure 1a), or in \( f_j(t) \) (Figure 1b) provides exactly the desired log-spaced support along the frequency axis. Subtracting neighboring cells is equivalent to taking a first order derivative with respect to the cell index, which amends to using the operator \( sL^{-1}_{k+1} \):

\[
\tilde{f}_s(\omega) = \frac{1}{(1 + \frac{i\omega}{s_j})^{k+1}} - \frac{1}{(1 + \frac{i\omega}{C s_j})^{k+1}}.
\]

where when integer \( k = 0 \) we are using \( F_s(\omega) \) scaled by \( s \), while when \( k > 0 \) we are using \( \tilde{f}_j(t) \). Integer \( j \) goes from 1 to \( N - 1 \). \( \omega \) is given as a function of \( s \) and it represents an internal frequency-line, similarly as \( \tilde{t} \) represents an internal timeline. The relationship between \( \omega \) and \( s \) will be explicitly derived later. Notice that \( s \) and \( \omega \) always appear together as ratio \( \frac{\omega}{s} \). This indicates the scale invariance since if \( \omega \) changes, scaling \( s \) by the same amount will result in a same transfer function.

Because \( \tilde{f}_s(\omega) \) is a result of subtraction of the activity of two low-pass filters its frequency characteristics will be band-pass. The central frequencies of the band-pass filters, \( \tilde{\omega} \), are the peaks of \( |\tilde{f}_s(\omega)| \). They can be analytically expressed in a simple form only for \( k = 0 \):

\[
\frac{\partial |\tilde{f}_s(\omega)|}{\partial \omega} = 0 \Rightarrow \tilde{\omega} = s \sqrt{C}.
\]

Inserting \( \tilde{\omega} \) back into equation (5) gives us the amplification of the band-pass filters: \( \frac{C^{-1} - 1}{C + 1} \). Notice that the amplification does not depend on \( s \) so it will be the same for all band-pass filters.

To identify the distribution of the central frequencies along the frequency axis we observe the ratio of the neighboring central frequencies:

\[
\frac{\tilde{\omega}_j}{\tilde{\omega}_{j+1}} = \frac{1}{C}.
\]

The constant ratio that is smaller than 1 (since \( C > 1 \)) indicates that the central frequency grows exponentially, providing log-spaced temporal receptive fields (Figure 2c and 2d). These ordered temporal receptive windows allow us not only to sort the signal by the speed of the components, but also to have an estimate of the speed.

We discussed above the case \( k = 0 \), notice that for any integer \( k \geq 0 \) the ratio \( \frac{\tilde{\omega}_j}{\tilde{\omega}_{j+1}} \) is constant. Increasing \( k \) increases the order of the low-pass filters and consequently makes the resulting band-pass filters more steep.

3. DEMONSTRATION

In this section we demonstrate how the proposed method for frequency decomposition, combined with PCA, can be used for identification of spatiotemporal features of interest. Given a multidimensional input, the frequency decomposition described above should be done independently for each dimension. One can then apply PCA across all the input dimensions separately for each band (\( N - 1 \) PCAs, where \( N \) is the number of time constants \( s \)) to quantify the energy of dominant components. Projecting the input signal on the eigenvectors that correspond to the largest eigenvalues (across bands) will give us the spatiotemporal components of interest. Additionally, directions of the eigenvectors will indicate which bands are representing the same components. Notice that since PCA is applied in each band separately we will be able to quantify the speed of the identified components.

We constructed a two-dimensional signal such that each dimension was composed of a weighted sum of a slow 7 Hz sine wave \( (x_s) \) and a fast 71 Hz cosine wave \( (x_f) \), Figure 3a. To each input we added binary multiplicative noise designed as a square wave with a duty cycle randomly varying between 1 and 10 ms. This was introduced to mimic occlusions that are generally present in visual and other types of input. Total duration of the signal was 5s. Our goal here was to isolate the components \( x_s \) and \( x_f \) assuming that we do not have any prior knowledge about the input signal.

The presence of multiplicative noise made the problem not well suited for SFA. SFA relies on taking a time-derivative of the input data to get an estimate of the velocity. Numerical derivative of the given input signal resulted in a rather noisy output that obscures the true components. The output of a classical SFA with 0.1 ms lagged time-derivative is shown in the first column in Figure 3b. Assuming we know the scale of the target signal we can reduce the noise by analyzing the signal at that scale. To illustrate this we filtered the input signal with a first order low-pass filter with a cutoff frequency of 20 Hz (different choices of the cutoff frequency led to similar or worse results). The result of applying SFA on such signal is shown in the second column in Figure 3b. The slow, 7 Hz component was estimated fairly well in this case, while the fast, 71 Hz component was still not accessible.

Since in general we do not know the target frequencies a priori it is not possible to do the filtering at the optimal scale. This is where the proposed method becomes useful: we can exploit the fact that we have a memory of the recent past and decompose the signal into multiple bands spanning all the frequencies of interest (from some minimum to some maximum frequency). We constructed a scale-invariant memory of each of the two input dimensions with \( N = 20 \) and \( k = 0 \). From this memory representation we built a frequency decomposi-
We proposed a biologically plausible neural method for constructing a scale-invariant frequency decomposition through a set of band-pass filters with log-spaced central frequencies. We demonstrated that this frequency decomposition has properties analogous to those of an internal timeline that can be constructed from the same memory representation. We characterized these properties in terms of the central frequency and gain and described how the frequency decomposition can be constructed from neurally plausible leaky integrators.

The crucial properties of the proposed method are the scale-invariance and log-spacing. When designing neural systems for processing natural signals it is essential to account for processing on multiple scales that span from milliseconds to hundreds of seconds, since the natural signals tend to be autocorrelated on a variety of temporal scales. Maintaining the same frequency resolution for all the frequency bands would require the number of bands to grow linearly with the frequency. This would be rather costly in terms of resources and computationally inefficient since we generally do not need high resolution at high frequencies. Therefore the proposed method provides a good trade-off between resolution and resources, in fact an optimal trade-off when the input signals are scale-invariant. Notice that memory systems based on register-style buffers or shift-registers could not account for the desired gradually decaying temporal resolution. The parallel design in our approach (leaky integrators operating in parallel, independently of each other) provides a simple solution for gradual decay in temporal accuracy which then allows us to construct the frequency-line with the desired properties.

Ability to decompose the input signal into multiple frequency bands can be particularly useful for extracting features that change slowly relative to the sensory representation. For instance, the activity of visual receptors changes rapidly comparing to the identity or position of the object in front of the eye. Existing unsupervised approaches such as SFA require some prior knowledge about the temporal scale of the signal. Additionally, band-pass decomposition allows not only identification of the components of interests, but also an estimate of their speed on a log-scale.

The proposed method does not require convolution like many other frequency decomposition methods. This makes it particularly suitable for neural implementation. Since our approach is based on biologically plausible neurons, neuromorphic platforms (see [22] for a recent example) could benefit for this approach. Designing neurons that can have long time constants would be a step towards making these computations simple to implement in neuromorphic hardware. Even though long time constants present a large challenge in analog electronics, the simplicity of their neural implementation [17] could provide an inspiration. In biological systems on a single-neuron level diffusion time constants are of the order
of 1 s, but when combined with renewal mechanisms they can be extended to infinity (demonstrated by frequently observed persistent firing). Having such long time constants the proposed method only requires lateral inhibition mechanisms for taking spatial derivatives.

5. REFERENCES


