

# Wavelet Scattering Transform for Variability Reduction in Cortical Potentials Evoked by Pitch Matched Electro-acoustic Stimulation in Unilateral Cochlear Implant Patients

Mehrdad Heydarzadeh, Sara Akbarzadeh, Chin-Tuan Tan

Department of Electrical and Computer Engineering, The University of Texas at Dallas, Richardson, USA

{mehhey, Sara.Akbarzadeh, Chin-Tuan.Tan}@utdallas.edu

**Abstract**—Cochlear implant (CI) restores the hearing sensation in profoundly deafen patients by directly stimulating auditory nerve with electric pulses using an array of tonotopically inserted electrodes. Basal electrodes stimulate in response to high input frequencies while apical electrodes stimulate to low input frequencies. The problem with this electrical stimulation, particularly in unilaterally implanted users who has residual hearing in the contra-lateral ear, lies in the frequency mismatch between characteristic frequency of auditory nerve and input signal. In this paper, we revisit our previously proposed mechanism for tuning intra-cochlear electrode to its pitch matched frequency using a single channel EEG [1]. We apply the wavelet scattering transform to extract a deformation invariant from the EEG signal recorded from each of 10 CI subjects when they were listening to pitch matched electro-acoustic stimulation. Results show that the wavelet scattering transform is able to capture the variability introduced by different subjects, and a more robust alternative to reveal the underlying neuro-physiological responses to this perceptual event.

**Index Terms**—Analytical complex wavelet, Cochlear implants, wavelet scattering transform

## I. INTRODUCTION

### A. Motivation

Cochlear Implant (CI) is a surgically implanted electronic device that provides a sense of sound to a patient who is profoundly deaf. It consists of a microphone and a speech processor that reside outside the head, which process input speech signal and encoded with electrical pulses. The information of electrical stimulation is transmitted to an array of electrodes implanted in the cochlea via a radio-frequency (RF) channel. Fig. 1 shows a typical cochlear implant. The intra-cochlear electrodes decode the signal and stimulate the auditory nerve tonotopically bypassing the peripheral auditory system. [2].

Direct electrical stimulation to cochlear nerve does not necessary match up the input acoustic frequency with the characteristic frequency of the cochlear nerve at where the intra-cochlear electrode is placed. As a result, individualized tuning of appropriate speech processing and electrical stimulation parameters is necessary for CI users. To date, this tuning process is done with verbal responses from the CI users when they listening to speech with an audiologist. This process is time and energy consuming and might not be helpful for very young users. However, CI users will need time to adapt to their devices to resolve this impoverish

representation of speech provided by electrical stimulation to derive benefits. An automatic tuning and evaluation platform is definitely a useful tool to CI users. With the advancement in neuro-engineering, a closed loop CI tuning system [3] and a signal processing algorithm for reducing the mismatch using independent component analysis [4] were also proposed in attempt to resolve this issue.

Cortical response to pitch matched stimuli was found to be a useful feedback for this tuning process. In [3], unilateral CI users with residual hearing in their non implanted ear were studied behaviorally and physiologically in matching electro-acoustic stimulation. CI users were asked to match the pitch percept elicited by electrical stimulation with those elicited by adjustable acoustic tone presented in the contra-lateral ear. EEG was also recorded when pitch matched stimuli were presented to these CI users.

One difficulty in interpreting the EEG signal is the low signal to noise ratio which hinders the extraction of the underlying physiological response and raises difficulty to establish a causal or statistical relationship between the behavioral and neural activities. Considering a zero mean additive noise for the measured signal, a typical method in reducing noise is ensemble averaging the waveform collected from several experiments. Since noise is zero mean, the ensemble average contains just the deterministic part of the signal. However, the ensemble averaging corrupts the deterministic part of this signal since the deterministic part the signal is not fixed waveform and is a time varying phenomena. Different factors like artifacts and other neural activities deform the deterministic part. As a result, the measured signal has a different shape. In our CI application, it is required to reduce the existing variability in order to establish a meaningful relation between EEG and the frequency mismatch.

### B. Key Contribution

In this paper, we consider a time warp model for deformation in EEG signal. Then, we apply the wavelet scattering transform to a single channel EEG before ensemble averaging which extracts a time warp deformation invariant from the signal. We applied this method to the data collected from 10 CI subjects. The experimental results show similarity between different subjects in the scattering domain.

This paper is outlined as below: Section II reviews the deformation model, the analytical complex wavelet and the scattering transform. The experimental results are reported in Section III. Finally, the concluding remarks are in Section IV.

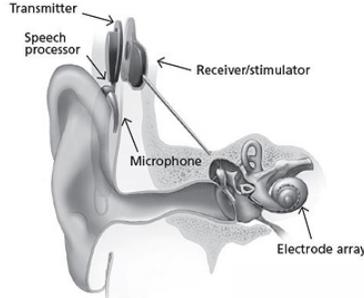


Fig. 1. A typical cochlear implant.

## II. SIGNAL PROCESSING FRAMEWORK

### A. Deformation and Notion of Stability

From a mathematical perspective, different EEG signals for the same neural activity are deformed versions of each others. Assuming a model for deformation can lead to developing robust signal processing methods. For example, consider the time shift deformation, which is defined as  $x(t) \rightarrow x_\tau(t - c)$  for an arbitrary signal  $x(t)$ . Then, the Fourier representation  $\hat{x}(\omega) = \int x(t)e^{-j\omega t} dt$  can be used for rejecting this deformation since  $|\hat{x}(\omega)|$  is shift invariant.

Time-warping is a general class of deformation which is modeled as  $x_\tau(t) = x(t - \tau(t))$  with  $|\tau'(t)| < 1$ . Here, the deformation can be quantified using  $|\tau'(t)|$ . If the derivative is zero the deformation simplifies to a pure shift. Suppose  $\Phi(x)$  is an arbitrary representation (feature) of  $x(t)$  and  $\Phi(x_\tau)$  is the same representation for  $x_\tau(t)$ , the deformed version of  $x(t)$ ; then,  $\Phi(\cdot)$  is considered as a robust representation if the difference between  $\Phi(x)$  and  $\Phi(x_\tau)$  caused by the deformation is small. This similarity can be quantified using the Euclidean norm as  $d(x, x_\tau) = \|\Phi(x) - \Phi(x_\tau)\|$ . Basically, we are looking for a  $\Phi(\cdot)$  which is invariant to this deformation. This invariance can be characterized using notion of stability.  $\Phi(\cdot)$  is stable if a small change in  $x(t)$  does not lead to a big change in  $d$ . Mathematically, stability is defined as Lipschitz continuity condition respect to the norm, if there exists a constant  $C > 0$  such that for all  $\tau$  with  $\sup_t |\tau'(t)| < 1$ :

$$\|\Phi(x) - \Phi(x_\tau)\| \leq C \sup_t |\tau'(t)| \|x\| \quad (1)$$

where  $\sup$  denotes supremum of a set. The constant  $C$  gives a measure of stability. Indeed, if the Lipschitz condition holds true, a change in  $x(t)$  caused by time warp leads to a linear change in its representation. Since, time-warping is locally linearized by  $\Phi(x(t))$  and  $\Phi(x) - \Phi(x_\tau)$  can be approximated by a linear operator if  $\sup_t |\tau'(t)|$  is small. In other words, in the feature space  $\Phi(x)$  and  $\Phi(x_\tau)$  are on the same hyperplane and the corresponding features do not spread all over the space.

### B. Analytic Wavelet Transform Modulus

The analytic wavelet transform, which is robust against shift deformation, can be calculated using constant  $Q$  filterbanks [6]. A wavelet like  $\psi(t)$  is a band pass filter where  $\hat{\psi}(0) = 0$ . In the analytical wavelet transform  $\hat{\psi}(\omega) \simeq 0$  for  $\omega < 0$  [7]. A dilated version of  $\psi(t)$  with the central frequency of  $\lambda > 0$  can be written as  $\psi_\lambda(t) = \lambda\psi(\lambda t)$  or in frequency domain  $\hat{\psi}_\lambda(\omega) = \hat{\psi}(\frac{\omega}{\lambda})$ , where the central frequency of  $\hat{\psi}(\omega)$  is normalized to 1 and  $Q$  is chosen as the number of wavelets per octave,  $\lambda = 2^{k/Q}$  for  $k \in \mathbb{Z}$ , which guarantees the bandwidth of  $\hat{\psi}_\lambda$  to be in order of  $Q^{-1}$  and its central frequency is at  $\lambda$ . In this way, different  $\hat{\psi}_\lambda$ 's cover all frequency axis except DC which is covered using a low-pass filter  $\phi$ . Let  $\Lambda$  denote the set of all values of  $\lambda$ , the wavelet transform of signal  $x(t)$  can be calculated by convolution of these filters:

$$Wx = (x(t) * \phi(t), x * \psi_\lambda(t)) \quad t \in \mathbb{R}, \lambda \in \Lambda \quad (2)$$

Here,  $t$  is not critically sampled as the wavelet bases, so this representation is redundant. The filters  $\phi$  and  $\psi$  need to be designed such that the entire frequency axis is covered which requires:

$$A(\omega) = |\hat{\phi}(\omega)|^2 + \frac{1}{2} \sum_{\lambda \in \Lambda} (|\hat{\psi}_\lambda(\omega)|^2 + |\hat{\psi}_\lambda(-\omega)|^2) \quad (3)$$

for all  $\omega \in \mathbb{R}$  satisfies [8]:

$$1 - \alpha \leq A(\omega) \leq 1 \quad \text{for } \alpha \leq 1 \quad (4)$$

By multiplying both sides of this inequality by  $|\hat{x}(\omega)|^2$  and apply Plancherel theorem one can obtain [9]:

$$(1 - \alpha) \|x\|^2 \leq \|Wx\|^2 \leq \|x\|^2 \quad (5)$$

where  $\|Wx\|^2 = \int |x * \phi|^2 + \sum_{\lambda \in \Lambda} \int |x * \psi|^2$  is the squared norm of wavelet representation and  $\|x\|^2 = \int |x(t)|^2 dt$  is the norm of signal. In Eq. 5, the lower bound guarantees a stable inverse while the upper bound shows that wavelet is a contractive operator [10]. If  $\alpha = 0$ , then  $W$  becomes a tight frame and  $x(t)$  can be reconstructed as  $x(t) = (x * \phi(t)) * \phi(-t) + \sum_{\lambda \in \Lambda} \text{Real}\{(x * \psi(t)) * \psi(-t)\}$  [8].

In the scattering transform, the wavelet modulus is used for feature extraction. In spite of Fourier transform, which is not possible to reconstruct the signal just using its Fourier modulus, it is possible to reconstruct the signal using just modulus of complex wavelet [11]. This is due to the redundant representation in Eq. 2. In addition, since the complex modulus is contractive,  $\|a\| - \|b\|$  for any  $(a, b) \in \mathbb{C}$ , the wavelet modulus operator,  $|W|$  is contractive:

$$\||W|x - |W|x'\|^2 \leq \|Wx - Wx'\|^2 \leq \|x - x'\|^2 \quad (6)$$

### C. Wavelet Scattering Transform

The main idea behind the scattering transform is to analyze the signal using analytical wavelet and then average the wavelet coefficients over time to extract features. The intuition behind the averaging coefficient is to reduce the variability in features and is similar to averaging the Fourier coefficient over

Mel frequency intervals to extract Mel frequency cepstral coefficients (MFCC) in speech processing [12]. However, MFCC loses information by averaging, while scattering transform preserves the reconstruction information [13].

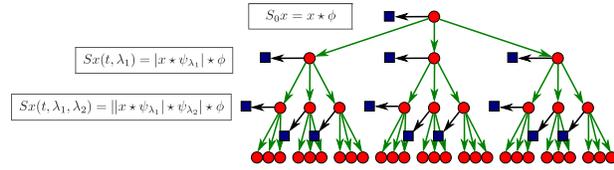


Fig. 2. Scattering transform of signal  $x$  by iterating the  $|W|_m$  operator. The output nodes are shown as squares.

In the scattering transform, a locally translation invariant descriptor is obtained by a time average  $S_0x(t) = x * \phi(t)$  which removes the high frequency contents. However, these high frequency content are recovered by a wavelet modulus transform as  $|W|_1x = (x * \phi, |x * \psi_{\lambda_1}|)$ . The first order of the scattering coefficients can be obtained as  $S_1x(t, \lambda_1) = |x * \psi_{\lambda_1}| * \phi$ . These coefficients measure the average signal amplitude in the frequency interval covered by  $\psi_{\lambda_1}$  with bandwidth corresponding to  $Q_1$ . In essence, they are calculated by a second wavelet modulus operator as  $|W|_2|x * \psi_{\lambda_1}| = (|x * \psi_{\lambda_1}| * \phi, ||x * \psi_{\lambda_1}| * \psi_{\lambda_2}|)$ . So, the second order scattering coefficients are  $S_2x = ||x * \psi_{\lambda_1}| * \psi_{\lambda_2}|$  which are computed by a  $\psi_{\lambda_2}$  with a bandwidth corresponding to  $Q_2$ . Iterating this process defines the scattering coefficients at any desired order.

For any  $m \geq 1$ , iterated wavelet modulus convolutions are written as  $U_m(t, \lambda_1, \dots, \lambda_m)x = |||x * \psi_{\lambda_1}| * \dots * \psi_{\lambda_m}|$  where the  $m^{th}$  order wavelet have an octave resolution of  $Q_m$  and they satisfy condition in Eq. 4. Next; the  $m^{th}$  order scattering coefficients are obtained by averaging  $U_mx$  with  $\phi$  as  $S_mx(t, \lambda_1, \dots, \lambda_m) = U_mx(t, \lambda_1, \dots, \lambda_m) * \phi$ . So, the scattering decomposition of a signal with the maximum order of  $l$  is an iterative operation by applying  $|W|_{m+1}$  on  $U_mx$  to obtain  $S_mx$  and  $U_{m+1}x$  for  $0 \leq m \leq l$  where  $U_0x = x$ . The scattering transform is the collection of all coefficients from each order  $Sx = \{S_mx | 0 \leq m \leq l\}$ . Fig. 2 shows the decomposition graph.

One can prove that the scattering transform has the following properties [9]:

- Time warp deformation stability: it satisfies the Lipschitz condition, ( i.e., there exist a constant  $C$  for any  $x$  such that  $\|Sx - Sx_\tau\| \leq C \sup_t |\tau'(t)| \|x\|$ ).
- Contraction: since scattering is calculated by wavelet modulus, it is a contractive transform, (i.e.,  $\|Sx - Sx'\| \leq \|x - x'\|$ ). As a result of this property, the scattering transform is robust against the additive noise.
- Energy conservation: if the chosen wavelet is a tight frame, then the scattering transform preserves the norm, (i.e.,  $\|x\|^2 = \|Sx\|^2 + \|U_{l+1}x\|^2$ ). As a result,  $\|U_{l+1}x\|^2$  vanishes as  $l \rightarrow \infty$ . In practice, the coefficients become very small after a few iterations.

### III. EXPERIMENTAL RESULTS

#### A. Data Description

We developed a real-time pitch matching platform that stimulates the electrode of interest directly at the user’s comfortable level (MCL) via the NIC research interface provided by Cochlear Corporation [Sydney, Australia]. The acoustic tone is amplified by a prescribed NAL-R gain based on the subject’s pure tone audiogram in the un-implanted ear. The acoustic signal is further amplified by a Graham Slee Solo SRGII amplifier and presented via an ER-2 insert earphone.

All acoustic and electric signals are presented in an alternating sequence, in which electrical pulses are presented via stimulation of a single channel in the implanted ear for 500 ms followed by 500 ms of electrical inactivity, the acoustic tone is shaped by trapezoidal window with a rising/falling time of 10 ms to prevent spectral splatter.

In these experiments a single EEG channel placed at the center of scalp (electrode 20) is used to collect data from 10 subjects at the rate of 500 Hz. The stimuli is a pure tone varied for 6 different frequencies: 250, 375, 500, 625, 1,000 Hz. These frequencies are chosen to be one octave higher or lower than the target frequency of 500 Hz and at the center of apical or basal electrodes.

#### B. Variability Reduction

We apply a three stage algorithm for reducing variability of EEG signal. Figure 3 depicts the flowchart of the algorithm.

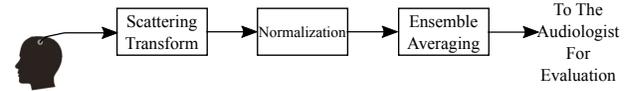


Fig. 3. A three step algorithm for variability reduction using the wavelet scattering transform.

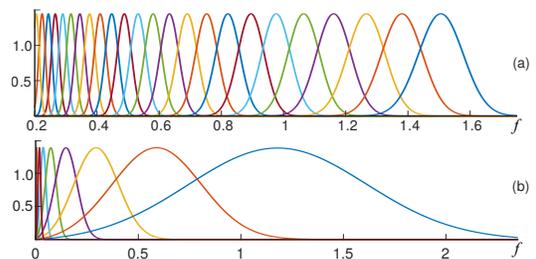


Fig. 4. Frequency response of Morlet wavelet used in scattering network. (a) First layer with  $Q_1 = 8$ , (b) Second layer with  $Q_2 = 1$ .

In the first step, the single channel EEG signal is analyzed with the scattering transform. We use a two layer ( $l = 2$ ) scattering network with the normalized Morlet wavelet defined as  $\psi(t) = e^{-i\omega_0 t} \theta(t) = e^{-it} e^{-t^2/2\sigma^2}$ , which is simply a modulated Gaussian function. In the frequency domain,  $\hat{\psi}(\omega) = \hat{\theta}(\omega - 1)$  is a low-pass filter with a Gaussian shape with its central frequency at the normalized frequency of 1. The Morlet wavelet is almost an analytical function since  $|\hat{\psi}(\omega)|$

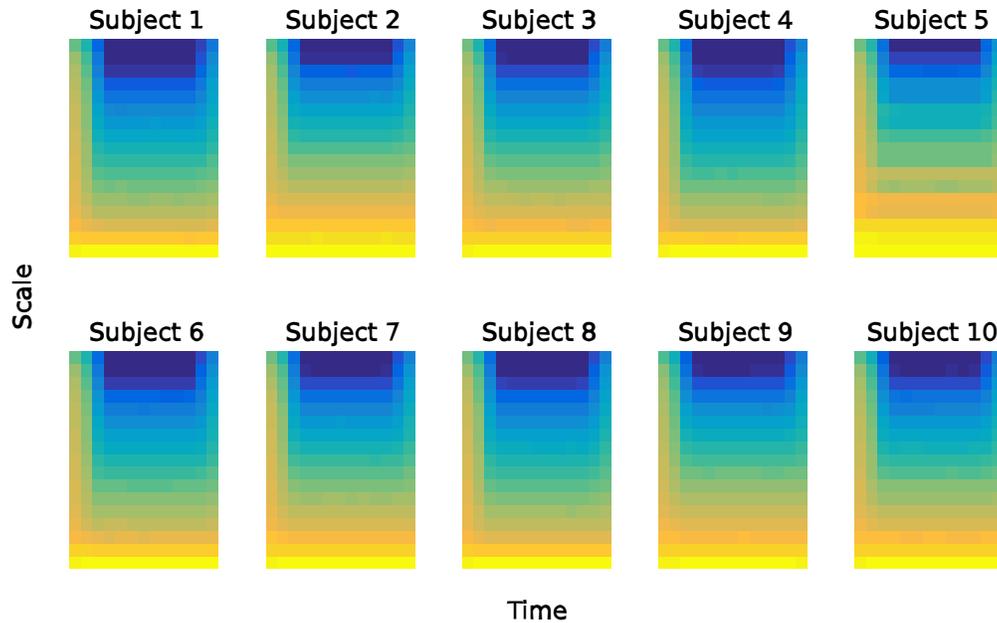


Fig. 5. Ensemble average of normalized scattering of EEG signal for matched pitch for 10 CI subjects. Different subjects show a high degree of similarity after removing the variability.

is small for  $\omega < 0$  but not zero. However, strictly speaking, Morlet is not an admissible wavelet [10]. For satisfying the admissibility condition we use  $\hat{\psi} = \hat{\theta}(\omega-1) - \hat{\theta}(\omega)\hat{\theta}(-1)/\hat{\theta}(0)$  which guarantees  $\hat{\psi}(0) = 0$ . The parameter  $\sigma^2$  determines the bandwidth of the wavelet which is assigned based on the choice of  $Q$  in scattering network. Fig. 4 shows the Morlet wavelets for  $Q_1 = 8$  and  $Q_2 = 1$  in our scattering network (chosen empirically). In the second stage, the scattering coefficients are normalized to be invariant against a change in the amplitude of input signal by dividing the coefficients in each layer by corresponding coefficients in the predecessor layer where the first layer coefficients are normalized by  $S_0x$ . After normalization, the log function is applied to the normalized coefficients as a range compressor. The range compression provides a better visualization for expert. In the last step, we apply an ensemble averaging to reduce noise.

Figure 5 depicts the ensemble average of the scattering transform for 10 CI subjects. As seen, the algorithm output shows a high degree of similarity for different subjects.

#### IV. CONCLUSION

In this paper, a method for tuning of cochlear implant based on EEG signal was investigated and the wavelet scattering transform was found to remove the variability in EEG signal. This processing framework provides a more robust alternative to the classical technique in mining the neurophysiological correlate to pitch matched stimuli, particularly in unilateral CI users with residual hearing in their contralateral non-implanted ear.

#### REFERENCES

- [1] C.-T. Tan, "A physiological correlate of electroacoustic pitch matching in cochlear implant users with residual hearing," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 5858–5862.
- [2] J. P. Roche and M. R. Hansen, "On the horizon: cochlear implant technology," *Otolaryngologic Clinics of North America*, vol. 48, no. 6, pp. 1097–1116, 2015.
- [3] M. Mc Laughlin, T. Lu, A. Dimitrijevic, and F.-G. Zeng, "Towards a closed-loop cochlear implant system: Application of embedded monitoring of peripheral and central neural activity," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 443–454, 2012.
- [4] M. Mc Laughlin, A. L. Valdes, R. B. Reilly, and F.-G. Zeng, "Cochlear implant artifact attenuation in late auditory evoked potentials: A single channel approach," *Hearing research*, vol. 302, pp. 84–95, 2013.
- [5] J. Birjandtalab, M. Heydarzadeh, and M. Nourani, "Automated eeg-based epileptic seizure detection using deep neural networks," in *Healthcare Informatics (ICHI), 2017 IEEE International Conference on*. IEEE, 2017, pp. 552–555.
- [6] K. N. Chaudhury and M. Unser, "On the shiftability of dual-tree complex wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 221–232, 2010.
- [7] I. Bayram and I. W. Selesnick, "A dual-tree rational-dilation complex wavelet transform," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 6251–6256, 2011.
- [8] S. G. Mallat, *A wavelet tour of signal processing the sparse way*. Amsterdam Boston: Elsevier/Academic Press, 2009.
- [9] S. Mallat, "Group invariant scattering," *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.
- [10] J. Andén and S. Mallat, "Deep scattering spectrum," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4114–4128, 2014.
- [11] S. Mallat and I. Waldspurger, "Phase retrieval for the cauchy wavelet transform," *Journal of Fourier Analysis and Applications*, vol. 21, no. 6, pp. 1251–1309, 2015.
- [12] J. R. Deller, J. H. Hansen, and J. G. Proakis, *Discrete-time processing of speech signals*. New York New York: Institute of Electrical and Electronics Engineers Wiley-Interscience, 2000.
- [13] J. Andén and S. Mallat, "Multiscale scattering for audio classification," in *ISMIR*, 2011, pp. 657–662.