DOMINANT COMPONENT TRACKING FOR EMPIRICAL MODE DECOMPOSITION USING A HIDDEN MARKOV MODEL

Steven Sandoval, Matthew Bredin, and Phillip L. De Leon

New Mexico State University
Klipsch School of Electrical and Computer Engineering
Las Cruces, New Mexico, U.S.A.
{spsandov,mbredin,pdeleon}@nmsu.edu

ABSTRACT

It is well known that the empirical mode decomposition algorithm does not always return an appropriate decomposition due to problems like mode mixing. In this paper, we consider the problem of a component being split across several intrinsic mode functions (IMFs). We propose the use of a hidden Markov model (HMM) to track the dominant component across the set of IMFs returned by EMD. We provide an example demonstrating the proposed tracking using an acoustic recording where component splitting is present in the decomposition and compare our method to two other possible tracking approaches. We show that the proposed method provides a compromise between smoothness and energy associated with the tracked component.

***Index Terms*—** Signal analysis, Empirical mode decomposition, Hidden Markov models

1. INTRODUCTION

In [1], Huang proposed the empirical mode decomposition (EMD) and sifting algorithms to sequentially determine a set of intrinsic mode functions (IMFs), $\{\varphi_k(t)\}$, where the definition of an IMF is any signal with real part that satisfies two conditions: 1) in the whole signal segment, the number of extrema and the number of zero crossings must be either equal or differ at most by one, and 2) at any point the mean value of the envelope, defined by the local maxima and the envelope defined by the local minima, is zero. IMFs are useful as signal components because they allow for a significant amount of flexibility in the signal model. However, IMFs are not orthogonal to each other and as a result, a decomposition into IMFs is not unique. Due to this ambiguity, the decomposition returned by EMD does not always capture the assumed/true underlying signal components as expected. More specifically, "mode mixing" and "component splitting," each of which will be further discussed in Section 2.1, may be present in the decomposition.

In this work, we assume that the signal under analysis contains a single dominant component which we wish to track across the IMFs returned by EMD. We propose a method for tracking the dominant component across the IMFs which utilizes a hidden Markov model (HMM) framework. The remainder of this paper is organized as follows. In Section 2 we briefly review the EMD and sifting algorithms as well as the mode mixing and component splitting problems. In Section 3 we review the HMM with time-varying transition probabilities. In Section 4 we describe our proposed approach. In Section 5 we provide simulations and results which illustrate the use of the proposed method. Finally, in Section 6 we provide concluding remarks.

2. EMPIRICAL MODE DECOMPOSITION

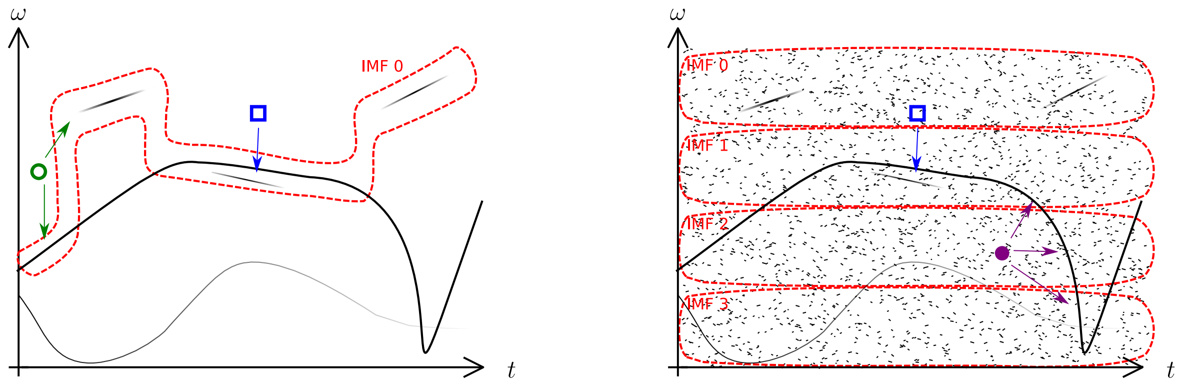
The EMD algorithm proposed by Huang [1] repeatedly calls the sifting algorithm, given in Algorithm 2. The purpose of the sifting algorithm is to iteratively identify and remove the trend from the signal, acting as a high pass filter. This process repeats to remove additional IMFs from the signal if they exist. The resulting decomposition is complete and sparse [1–3]. The sifting algorithm is an iterative way of removing the asymmetry between the upper and lower envelopes in order to transform the input $r(t)$ into an IMF [4].

Since the original publication, many improvements to EMD have been proposed to address computational issues and other issues related to the signal decomposition. The ensemble empirical mode decomposition (EEMD) [5] introduced ensemble averaging in order to address the mode mixing problem via an additive noise and an averaging of IMF estimates. The complete EEMD (CEEMD) was proposed to address some of the undesirable features of EEMD by averaging at the IMF level as each IMF is estimated rather than averaging at the conclusion of EEMD [6]. The improved CEEMD (ICEEMD) [7] was proposed to reduce the noise present in each IMF estimate and to reduce the occurrence of spurious IMFs as was observed with CEEMD. More recently, we proposed [8] additional improvements to CEEMD which include 1) a modification to the ensemble averaging which guarantees that the average IMF is a true IMF [5] and 2) a change from the additive noise used in ensemble averaging to a complimentary pair of narrowband tones which we term "tone masking." In this work, we utilize the decomposition algorithm proposed in [8]. For convenience, we have repeated Algorithms 1 - 3, however for complete details see [8].

2.1. Mode Mixing and Component Splitting

One major problem in the EMD algorithm is mode mixing, which is defined as a single IMF either consisting of components of disparate scales or components of similar scale residing in the same IMF [5]. Mode mixing is a consequence of signal intermittency, or more specifically relative component intermittency. As a result, the particular component(s) estimated by the sifting algorithm in a particular IMF at any instant may change as intermittent components begin or end [9]. This is illustrated in Fig. 1(a), where in the top part of the figure, we show components of disparate scales being in the same IMF denoted by ○, while in the center part of the figure we show two components of similar scale in the same IMF denoted by □. The ability of EMD to resolve two components considering both the relative instantaneous amplitudes (IAs) and instantaneous frequencies (IFs) of components, was examined and quantified by Rilling [10].

One commonly used method of mitigating mode mixing is



(a) (b)

Fig. 1. In (a) and (b), the assumed components are indicated with — and the first component or high frequency IMF, identified with the sifting algorithm, is indicated within the --- frame. In (a), the mode mixing problem is apparent where we see components of disparate scales being in the same IMF (indicated by ○) and components of similar scale in the same IMF (indicated by □). In (b), adding noise and ensemble averaging may assist in resolving mode mixing where components of disparate scales appear in the same IMF (○), however this may not help a situation where components of similar scale appear in the same IMF (□) and additionally component splitting may be a problem (indicated by ●).

EEMD [5]. EEMD utilizes zero-mean white noise to perturb the signal so a component may be extracted properly over an ensemble average. As the illustration in Fig. 1(b) shows, noise can be used to assist the sifting algorithm. Inserting noise with high enough power gives the sifting algorithm something to track when the highest frequency component is intermittent (indicated by ○ in Fig. 1(a)), then vanishes in the ensemble average. Although injecting noise can help to extract components properly, a carefully designed masking signal can also be used [8]. However, the use of masking signals may not help a situation where components of similar scale appear in the same IMF (indicated by □). The use of masking signals often results in a bandpass structured decomposition which may not have the flexibility to extract a component which has significant variation in frequency. This leads to the problem we term *component splitting* and is illustrated in Fig. 1(b). Ideally the component of interest (—) would be extracted as a single IMF. However, as is often the case, the component of interest with significant frequency variation may be "split" across several IMFs.

Algorithm 1 Decomposition and Demodulation [8]

procedure $\{\hat{\varphi}_k(t), \hat{a}_k(t), \hat{\omega}_k(t)\} = \mathrm{DecompDemod}(\ x(t)\ )$
  initialize: $x_{-1}(t) = x(t)$, $k = 0$, β_k is a SNR factor, ε_2 is an energy threshold, and I is the number of trials
  while $\int |x(t)|^2\,\mathrm{d}t / \int |x_{k-1}(t)|^2\,\mathrm{d}t > \varepsilon_2$ and $x_{k-1}(t)$ is not monotonic **do**
    $$\bar{\varphi}_k(t) = \frac{1}{2I}\sum_{i=1}^{I}\Big[\mathrm{SIFT}(\ x_{k-1}(t) + v^{(i,k)}(t)\ ) + \mathrm{SIFT}(\ x_{k-1}(t) - v^{(i,k)}(t)\ )\Big]$$
    $\hat{\varphi}_k(t) = \mathrm{SIFT}(\ \bar{\varphi}_k(t)\ )$
    $[\hat{a}_k(t), \hat{\omega}_k(t)] = \mathrm{IMFdemod}(\ \hat{\varphi}_k(t)\ )$
    $x_k(t) = x_{k-1}(t) - \hat{\varphi}_k(t)$
    $k \leftarrow k + 1$
  end while
  $\hat{\varphi}_k(t) = x_{k-1}(t)$
end procedure

Algorithm 2 Sifting Algorithm

procedure $\varphi(t) = \mathrm{SIFT}(\ r(t)\ )$
  while $\frac{1}{NT_s}\int_0^{NT_s} |\tilde{e}(t)|^2 dt \geq \varepsilon_1$ **do**
    find all local maxima: $u_p = r(t_p),\ p = 1, 2, \ldots$
    find all local minima: $l_q = r(t_q),\ q = 1, 2, \ldots$
    interpolate: $u(t) = \mathrm{CubicSpline}(\{t_p, u_p\})$
    interpolate: $l(t) = \mathrm{CubicSpline}(\{t_q, l_q\})$
    $e(t) = [u(t) + l(t)]/2$
    $r(t) \leftarrow r(t) - e(t)$
  end while
  $\varphi(t) = r(t)$
end procedure

Algorithm 3 IMF demodulation

procedure $[\hat{a}(t), \hat{\omega}(t)] = \mathrm{IMFdemod}(\ \hat{\varphi}(t)\ )$
  $\hat{a}(t) = \mathrm{IAest}(\ \hat{\varphi}(t)\ )$
  $\hat{s}_{\mathrm{FM}}(t) = \mathrm{iterAMremoval}(\ \hat{\varphi}(t)\ )$
  $$\hat{\sigma}_{\mathrm{FM}}(t) = -\mathrm{sgn}\left[\frac{d}{\mathrm{d}t}\hat{s}_{\mathrm{FM}}(t)\right]\sqrt{1^2 - \hat{s}^2_{\mathrm{FM}}(t)}$$
  Find $\{t_0\}$ such that $\hat{\sigma}_{\mathrm{FM}}(t_0) = 0$
  For each t_0, replace $(\hat{\sigma}_{\mathrm{FM}}(t_0 - \epsilon), \hat{\sigma}_{\mathrm{FM}}(t_0 + \epsilon))$ with interpolation
  $$\hat{\omega}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\arg\left[\hat{s}_{\mathrm{FM}}(t) + j\hat{\sigma}_{\mathrm{FM}}(t)\right]$$
end procedure

3. HIDDEN MARKOV MODELS

HMMs provide a probabilistic approach for relating a sequence of observations to a sequence of hidden classes or hidden states that explain the observations [11,12]. In this section, we introduce mathematical notion and review important computational algorithms related to HMMs that are utilized in Section 4 to describe the proposed component tracking.

Mathematically, the typical HMM consisting of K states $\mathbf{Q} = \{q_1\ q_2\ \ldots\ q_K\}$ with observation sequences $\mathbf{O}_N = [o_1\ o_2\ \cdots\ o_N]$

of length N may be specified by parameter set $\lambda = \{\mathbf{A},\ \mathbf{B},\ \mathbf{\Pi}\}$ where

- $\mathbf{A}$ is a transition probability matrix where $(\mathbf{A})_{ij} = \mathsf{a}_{ij},\ 1 \leq i \leq K,\ 1 \leq j \leq K$ representing the probability of moving from state i to state j
- $\mathbf{B}$ a matrix of observation likelihoods where $(\mathbf{B})_{kn} = b_k(o_n),\ 1 \leq k \leq K,\ 1 \leq n \leq N$ is the likelihood of observation o_n being generated from a state k at time n
- $\mathbf{\Pi} = [\pi_1\ \pi_2\ \ldots\ \pi_K]$ an initial probability distribution over states, i.e., π_k is the probability that the Markov model will start in state k.

However, in order to use the HMM framework for component tracking, the typical HMM must be generalized to allow for a time-varying transition matrix

$$(\mathbf{A})_{ij} = \mathsf{a}_{ij} \quad \rightarrow \quad (\mathbf{A}[n])_{ij} = \mathsf{a}_{ij}[n]. \tag{1}$$

Our analysis using HMMs will require the computation of three quantities

1. $\alpha_n[k]$, the probability of ending up at state q_k at time n, given observations $[o_1\ o_2\ \cdots\ o_n]$
2. $\beta_n[k]$, the probability of the observations $[o_{n+1}\ o_{n+2}\ \cdots\ o_N]$, given that we are in state q_k at time n
3. $q_n^\star$, the most likely state at time n

Fortunately there exist well-known and efficient algorithms for computing these quantities. More specifically, $\alpha_n[k]$, $\beta_n[k]$, and $q_n^\star$ may be respectively computed using the forward, backward, and Viterbi algorithms. The pseudocode for the forward, backward, and Viterbi algorithms generalized for a time-varying transition matrix are given in Algorithms 4-6.

Algorithm 4 Forward Algorithm

procedure $\alpha_n[k] = \text{forward}(\ \mathbf{A},\ \mathbf{B},\ \mathbf{\Pi}\ )$
  initialize: $\alpha_1[k] = \pi_k b_k(o_1),\ 1 \leq k \leq K$
  for $n = 2, 3, \cdots, N$ **do**
    for $k \in \{1, 2, \cdots, K\}$ **do**
      $\alpha_n[k] = \sum_{i=1}^{K} \alpha_{n-1}[i]\mathsf{a}_{ik}[n]b_k(o_n)$
    end for
  end for
end procedure

Algorithm 5 Backward Algorithm

procedure $\beta_n[k] = \text{backward}(\ \mathbf{A},\ \mathbf{B},\ \mathbf{\Pi}\ )$
  initialize: $\beta_N[k] = 1,\ 1 \leq k \leq K$
  for $n = N-1, N-2, \cdots, 1$ **do**
    for $k \in \{1, 2, \cdots, K\}$ **do**
      $\beta_n[k] = \sum_{j=1}^{K} \mathsf{a}_{kj}[n]b_j(o_{n+1})\beta_{n+1}[j]$
    end for
  end for
end procedure

Algorithm 6 Viterbi Algorithm

procedure $[q_n^\star, v^\star] = \text{Viterbi}(\ \mathbf{A},\ \mathbf{B},\ \mathbf{\Pi}\ )$
  initialize: $v_1[k] = \pi_k b_k(o_1),\ 1 \leq k \leq K$
  initialize: $p_1[k] = 0,\ 1 \leq k \leq K$
  for $n = 2, 3, \cdots, N$ **do**
    for $k \in \{1, 2, \cdots, K\}$ **do**
      $v_n[k] = \max_i\ v_{n-1}[i]\mathsf{a}_{ik}[n]b_k(o_n)$
      $p_n[k] = \operatorname{argmax}_i\ v_{n-1}[i]\mathsf{a}_{ik}[n]b_k(o_n)$
    end for
  end for
  $v^\star = \max_i\ v_N[i]$
  $q_N^\star = \operatorname{argmax}_i\ v_N[i]$
  for $n = N-1, N-2, \cdots, 2$ **do**
    $q_n^\star = p_n[q_{n+1}^\star]$
  end for
end procedure

4. PROPOSED TRACKING

At the conclusion of Algorithm 1 we have a set of K IMFs each with IA $\hat{a}_k(t)$ and IF $\hat{\omega}_k(t)$ parameters. As described in Section 2.1, EMD can suffer from mode mixing as well as component splitting. Suppose, we wish to isolate and track the dominant component across a set of IMFs in the presence of mode mixing and/or component splitting. Trivially, one could assume: 1) the dominant component is captured in the IMF with the highest total energy $\arg\max \int |a_k(t)|^2\, \mathrm{d}t\ \forall\ t$ or 2) the dominant component is simply the component associated with the IMF with the largest a_k, i.e. $\arg\max a_k(t)$ at every time instant. In the first case, the resulting dominant component estimate is smooth in frequency, while in the latter it may jump sporadically in frequency. On the other hand, the tracking for the second case would account for the maximum amount of energy possible.

Next, we propose an alternate approach which uses an HMM framework and allows for a compromise between the smoothness and energy capture of the component tracking. In order to encourage smoothness in the tracking, a time-varying transition matrix is constructed by considering how close the components are spaced to one another in frequency (IF) at subsequent instants in time as

$$\mathsf{a}_{ij}[n] = \frac{|\hat{\omega}_i(nT_s) - \hat{\omega}_j([n+1]T_s)|^{-1}}{\sum_{k=1}^{K} |\hat{\omega}_i(nT_s) - \hat{\omega}_k([n+1]T_s)|^{-1}}. \tag{2}$$

In order to encourage high energy in the tracking we build the observation likelihoods using the IAs of the components as

$$b_k(o_n) = \exp(|\hat{a}_k(nT_s)|) \tag{3}$$

and assume equally likely initial states

$$\pi_k = \frac{1}{K}. \tag{4}$$

Using the above, the forward and backward variables $\alpha_n[k]$ and $\beta_n[k]$ are computed and by Bayes' rule, the likelihood of a particular state at an instant in time is used as the observation likelihood

$$b_k(o_n) = \alpha_n[k]\beta_n[k]. \tag{5}$$

Finally, the Viterbi algorithm is used to determine the most likely sequence of states, i.e. the index of the dominant component at each instant in time. The pseudocode for dominant component tracking (DomComp) is given in Algorithm 7 and called after Algorithm 1.

5. EXAMPLES AND RESULTS

We choose to demonstrate the proposed dominant component tracking using a recording of a parakeet call which was made using a

Algorithm 7 Dominant Component Algorithm

procedure $\{a^\star[n], \omega^\star[n]\} = \text{DomComp}(\ \{\hat{a}_k(t), \hat{\omega}_k(t)\}\ )$

$\quad \mathsf{a}_{ij}[n] \leftarrow \frac{|\hat{\omega}_i(nT_s) - \hat{\omega}_j([n+1]T_s)|^{-1}}{\sum_{k=1}^{K} |\hat{\omega}_i(nT_s) - \hat{\omega}_k([n+1]T_s)|^{-1}}$

$\quad b_k(o_n) \leftarrow \exp(|\hat{a}_k(nT_s)|)$

$\quad \pi_k \leftarrow \frac{1}{K}$

$\quad \alpha_n[k] = \text{forward}(\ \mathbf{A},\ \mathbf{B},\ \mathbf{\Pi}\ )$

$\quad \beta_n[k] = \text{backward}(\ \mathbf{A},\ \mathbf{B},\ \mathbf{\Pi}\ )$

$\quad \beta_n[k] \leftarrow \alpha_n[k]\beta_n[k]$

$\quad [q_n^\star, v^\star] = \text{Viterbi}(\ \mathbf{A},\ \mathbf{B},\ \mathbf{\Pi}\ )$

$\quad a^\star[n] = a_{q_n^\star}(nT_s)$

$\quad \omega^\star[n] = \omega_{q_n^\star}(nT_s)$

end procedure

wireless "backpack" microphone positioned close to the bird's head. Recently, researchers have begun studying individual-level calling behaviour of animals that are behaving freely in social groups using wireless systems that consist of a microphone mounted in a wearable backpack and a receiver station [13–15]. These light-weight backpack systems typically consist of a small microphone, analog-to-digital converter, amplifier, microcontroller, FM transmitting chip, and battery.

The acoustic recording has sampling rate $f_s = 48$ kHz and was taken in a noisy social environment where much of the noise consisted of calls from other birds in environment. Therefore the target call is overlapped with noise in both time and frequency, however, the target call will typically have a greater amplitude than the noise due to the close proximity to the microphone. A short-time Fourier transform (STFT) for the recording is shown in Fig. 2(a) where we used a Hamming window of length $L = 256$ samples and a window advance of 4 samples. We use Algorithm 1 with masking signal

$$v^{(i,k)}(t) = \beta \sin\left[\frac{c^k}{2}t + \nu\right] \tag{6}$$

and $\beta = 3$, $c = 0.75$, $\varepsilon_1 = -120$ dB, $\varepsilon_2 = 5$ dB, and $I = 3$ trials. A time-frequency plot of the IA/IF parameters from the resulting eight components are shown in Fig. 2(b). We then apply Algorithm 7 to the IA/IF parameters to obtain an estimate of the dominant component shown in Fig. 2(c). This figure illustrates a single component which appears to trace out a smooth path of high energy. For example, in Fig. 2(b) near $t = 0.07$ we see one component spread across two IMFs with relatively large IA and in Fig. 2(c) the component has been properly tracked across IMFs. Similarly, in Fig. 2(b) near $t = 0.22$ we see one component spread across three IMFs and in Fig. 2(c) the component has been properly tracked across all three IMFs. For comparison, we show in Figs. 2(d) and (e) the results of the two trivial methods for dominant component tracking previously described in Section 4. In the former, we have selected the IMF based on highest total energy and in the latter, we have taken a path which used the maximum IA at every instant in time.

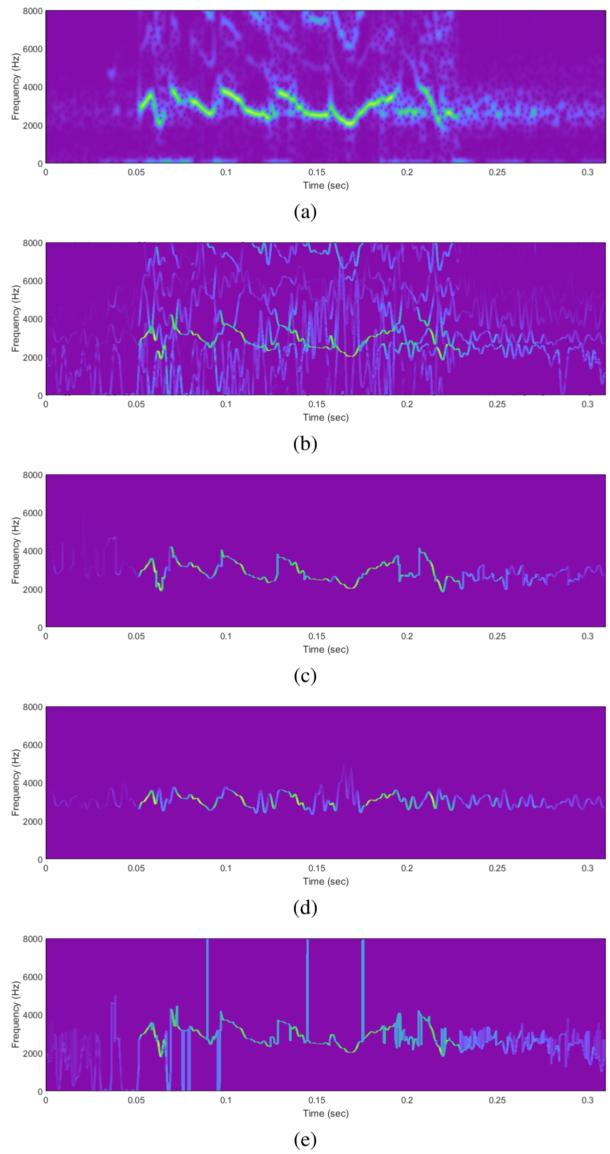


Fig. 2. Recording of a parakeet call and results from (a) STFT , (b) decomposition and demodulation (Algorithm 1), (c) proposed algorithm (Algorithm 7), (d) IMF with the greatest total energy, and (e) IMF at each time instant with the maximum instantaneous amplitude.

6. CONCLUSION

In this paper, we carefully described and considered the problem of component splitting in EMD. We proposed to use a HMM-based algorithm to track the dominant component across a set of IMFs. The transition probabilities are time-varying and based on how close the components are spaced in frequency. The observation likelihoods are based on the amplitudes of the components. We demonstrated the method using an acoustic recording where component splitting was present in the decomposition. We compared against two alternate methods (one where an IMF was selected based on highest total energy and another based on selecting the path which maximizes the IA at every instant in time) and showed that the proposed method led to a dominant component track which yields a compromise between smoothness and energy associated with the track.

7. ACKNOWLEDGEMENTS

The authors wish to thank Prof. Timothy Wright and Dr. Gregory Kohn of the Dept. of Biology at New Mexico State University for supplying the acoustic recording of a parakeet used in this research.

8. REFERENCES

[1] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. London Ser. A*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[2] T. Y. Hou and Z. Shi, "Adaptive data analysis via sparse time-frequency representation," *Adv. Adapt. Data Anal.*, vol. 3, no. 01n02, pp. 1–28, 2011.

[3] ——, "Data-driven time–frequency analysis," *Appl. Comput. Harmon. Anal.*, vol. 35, no. 2, pp. 284–308, 2013.

[4] R. Rato, M. Ortigueira, and A. Batista, "On the HHT, its problems and some solutions," *Mechanical Syst. Signal Process.*, vol. 22, no. 6, pp. 1374–1394, 2008.

[5] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 01, pp. 1–41, 2009.

[6] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2011, pp. 4144–4147.

[7] G. Colominas, Marcelo A. and Schlotthauera and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal. Process. Control*, vol. 14, pp. 19–29, Nov. 2014.

[8] S. Sandoval and P. L. De Leon, "Advances in empirical mode decomposition for computing instantaneous amplitudes and instantaneous frequencies," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2017, pp. 4311–4315.

[9] R. Deering and J. Kaiser, "The use of a masking signal to improve empirical mode decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2005, pp. 485–488.

[10] G. Rilling and P. Flandrin, "One or two frequencies? The empirical mode decomposition answers," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 85–95, 2008.

[11] L. R. Rabiner and B.-H. Juang, *Fundamentals of speech recognition*. PTR Prentice Hall Englewood Cliffs, 1993, vol. 14.

[12] D. Jurafsky and J. H. Martin, *Speech and Language Processing*. Pearson Prentice Hall, 2009.

[13] V. N. Anisimov, J. A. Herbst, A. N. Abramchuk, A. V. Latanov, R. H. Hahnloser, and A. L. Vyssotski, "Reconstruction of vocal interactions in a group of small songbirds," *Nature Methods*, vol. 11, no. 11, p. 1135, 2014.

[14] L. F. Gill, W. Goymann, A. Ter Maat, and M. Gahr, "Patterns of call communication between group-housed zebra finches change during the breeding cycle," *Elife*, vol. 4, 2015.

[15] D. Stowell, E. Benetos, and L. F. Gill, "On-bird sound recordings: Automatic acoustic recognition of activities and contexts," *IEEE Trans. Audio, Speech, and Lang. Process.*, vol. 25, no. 6, pp. 1193–1206, 2017.