

「局所的位相検出法」を用いた実測データの処理
--- DCASE2020-task2-fan ---

2020.11.22.(Sun.)
園部和夫

0. はじめに

以前に、信号源とセンサー（マイクロホン）の間に生じる予め測定できない不明な伝達特性をバイパスし、信号源内部の原信号を回復できる方法を考案し「局所的位相シフト検出法」と命名した。

URL : <https://sigprocrandwalk.hatenadiary.org/entry/20170722/1500751760>

URL : <https://sigprocrandwalk.hatenadiary.org/entry/2020/03/22/193447>

以上のリンク先では、信号源に「非線形性」が作用していて各帯域間が独立していない場合（相互変調の状態）に、観測信号で得られる局所的な位相の進み遅れ量を用いて原信号を回復している。

以下では、実測データである「DCASE2020」
（<http://dcase.community/challenge2020/index>）
用に公開されたデータのうち「task2」の「fan」に対して「局所的位相シフト検出法」を用いた場合を示す。

1. 「DCASE2020」の「task2」の「fan」のデータについて

「DCASE2020」の「task2」の「fan」に含まれる先頭の2つのデータに対して、「従来方法（ハニング窓による切り出し+フーリエ変換を用いたパワースペクトル表示）」によるスペクトログラムと、「局所的位相シフト検出法」によるスペクトログラムと、を求め結果を比較する。

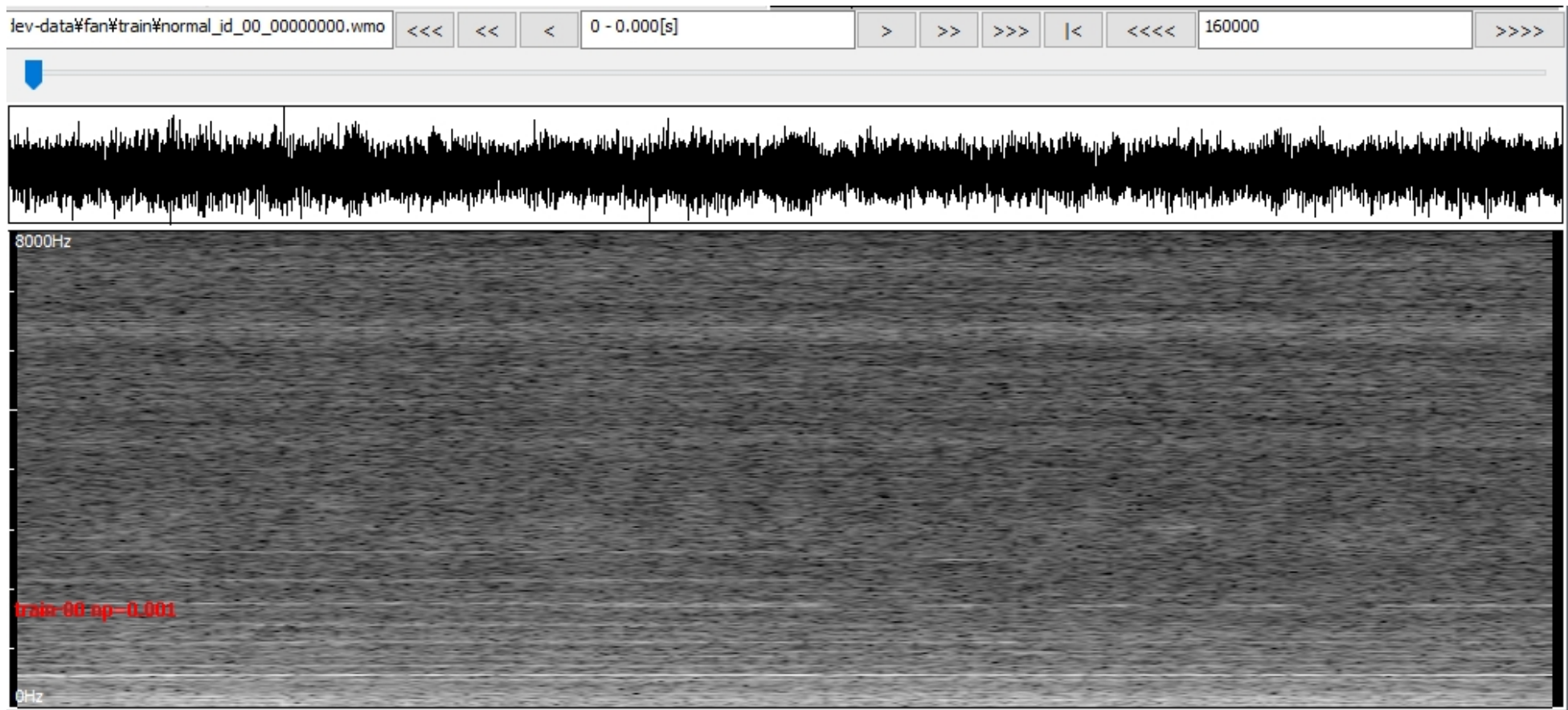


図 1-1-1. normal_id_00_00000000.wav に対する
「従来方法」によるスペクトログラム (0Hz-8000Hz)

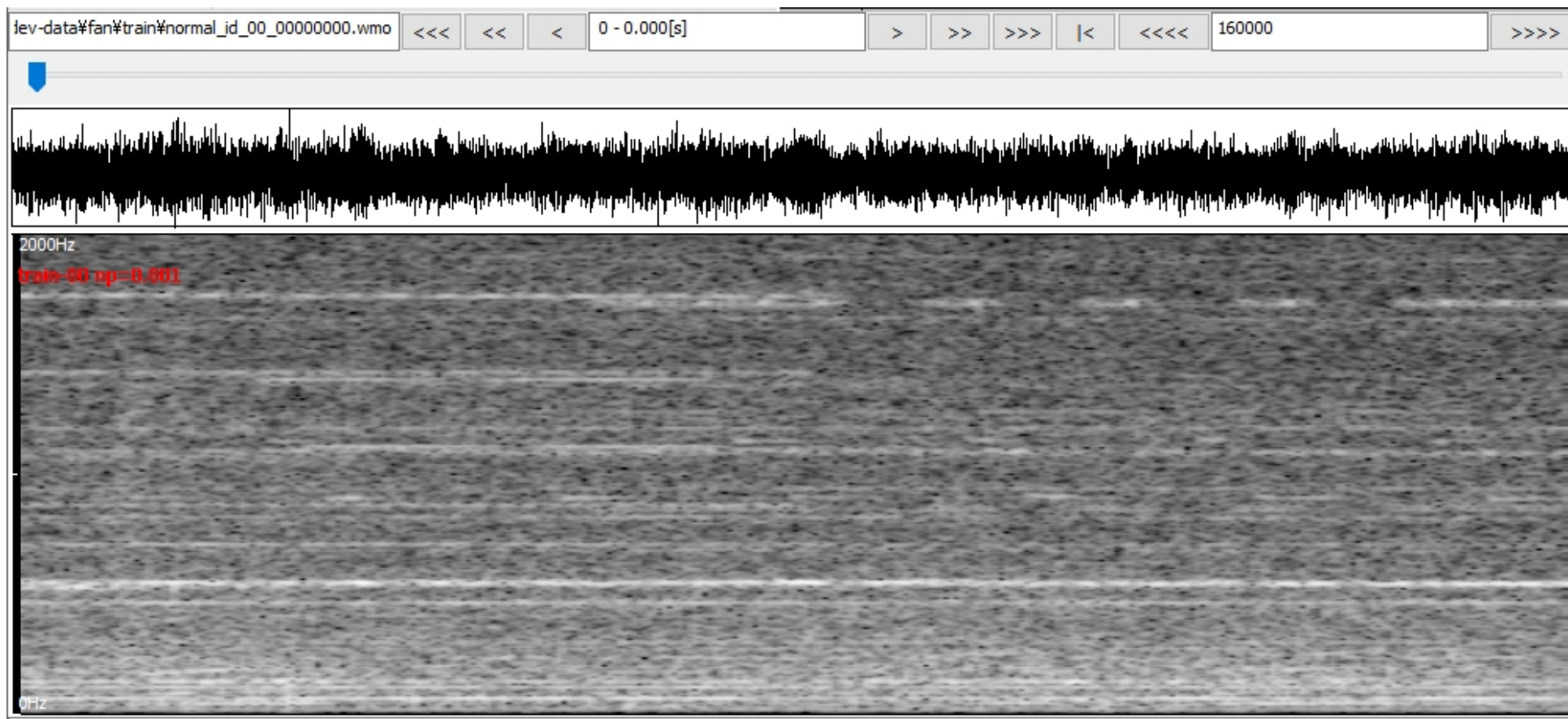


図 1-1-2. normal_id_00_00000000.wav に対する
「従来方法」によるスペクトログラム (0Hz-2000Hz)

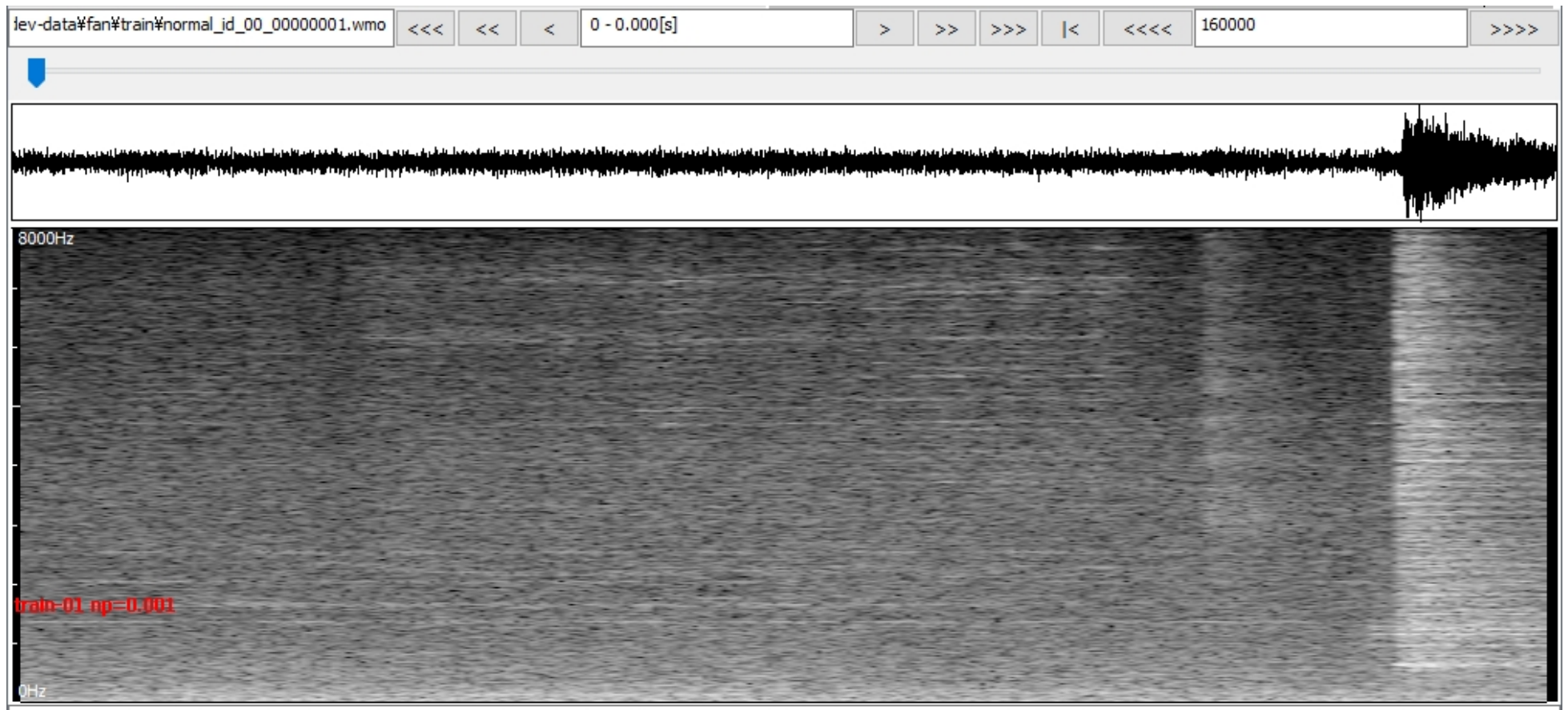


図 1-2-1. normal_id_00_00000001.wav に対する
「従来方法」によるスペクトログラム (0Hz-8000Hz)

データに付けられたファイル名（normal_id_00_00000000.wav など）のサフィックスは次の意味を表している。

左：normal（normal か anomaly の normal）

中：id_00（複数あるデバイス（fan）のデバイス番号）

右：00000000（同一デバイスに対して複数試行を行った際の番号）

以上のサフィックスより、図 1-1-x. と図 1-2-x. は同一デバイスにおける異なる試行番号の結果を表している。図 1-1-x. と図 1-2-x. を目視で見ると、「従来方法」によるスペクトログラムでは、同一デバイスにも拘わらず結果が安定して見えない。ここが安定していない場合、続けて処理（統計的分類や機械学習など）を行う場合に不利になる可能性がある。

以下では、図 1-1-x. と図 1-2-x. と同一のデータに対して、「局所的位相シフト検出法」を用いた場合のスペクトログラムを示す。

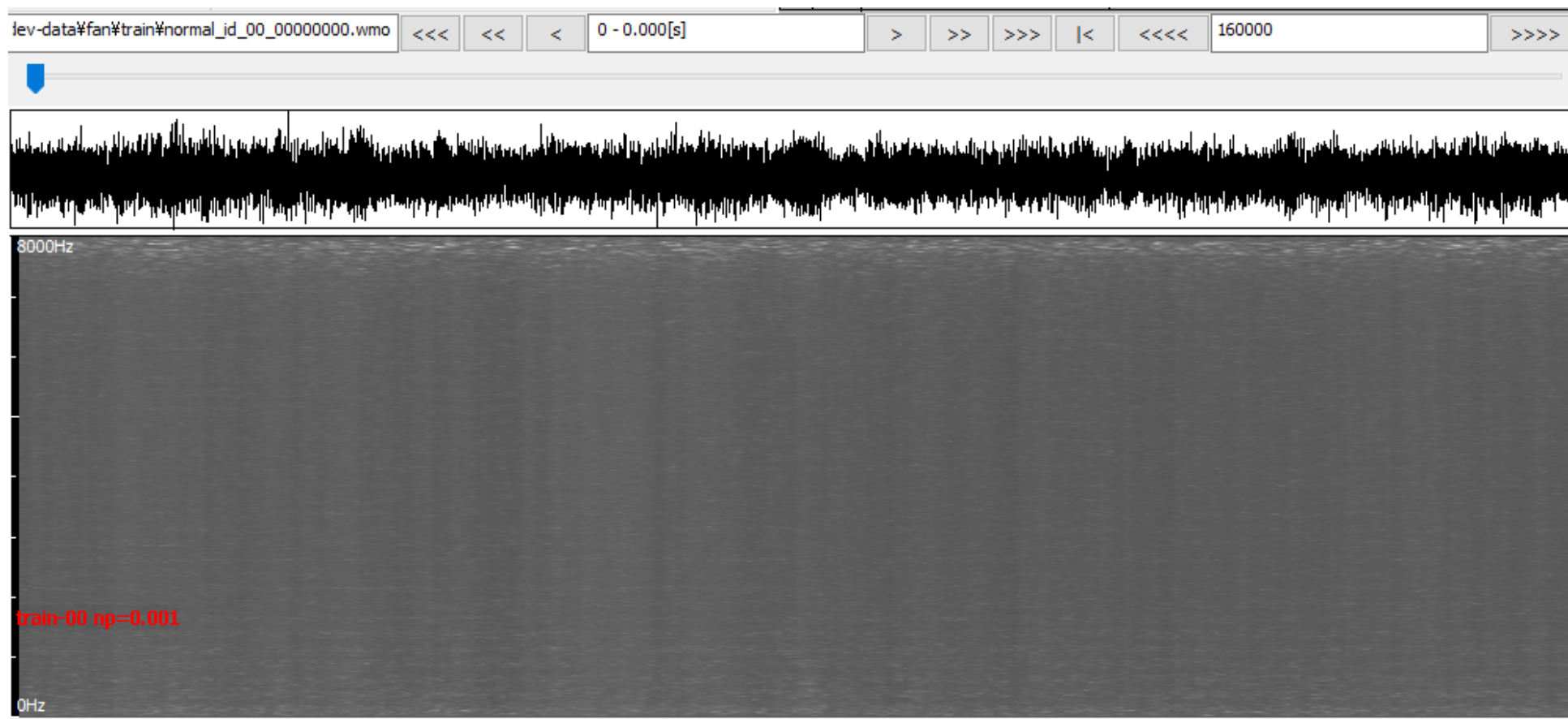


図 1-1-1-rst. normal_id_00_00000000.wav に対する
「局所的位相シフト検出法」によるスペクトログラム (0Hz-8000Hz)

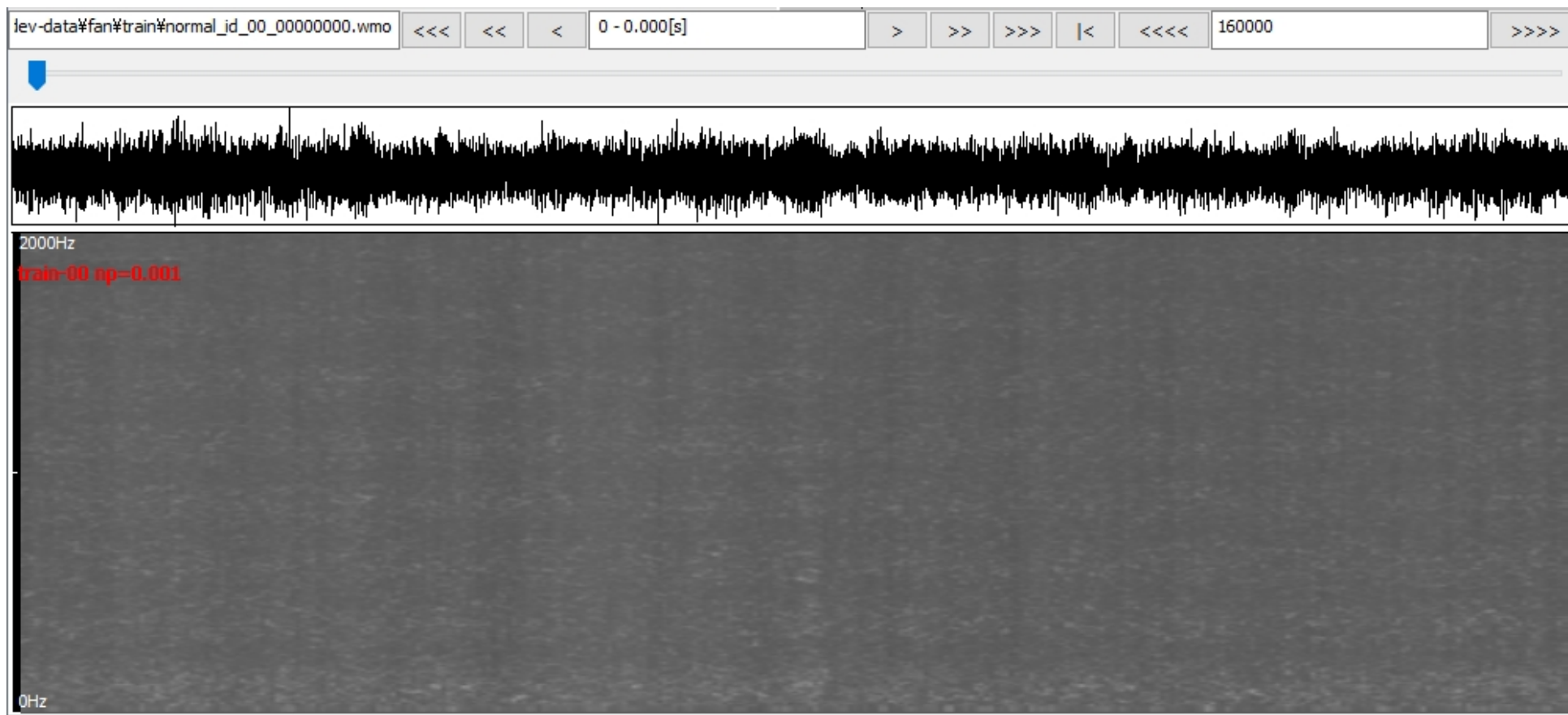


図 1-1-2-rst. normal_id_00_00000000.wav に対する
「局所的位相シフト検出法」によるスペクトログラム (0Hz-2000Hz)

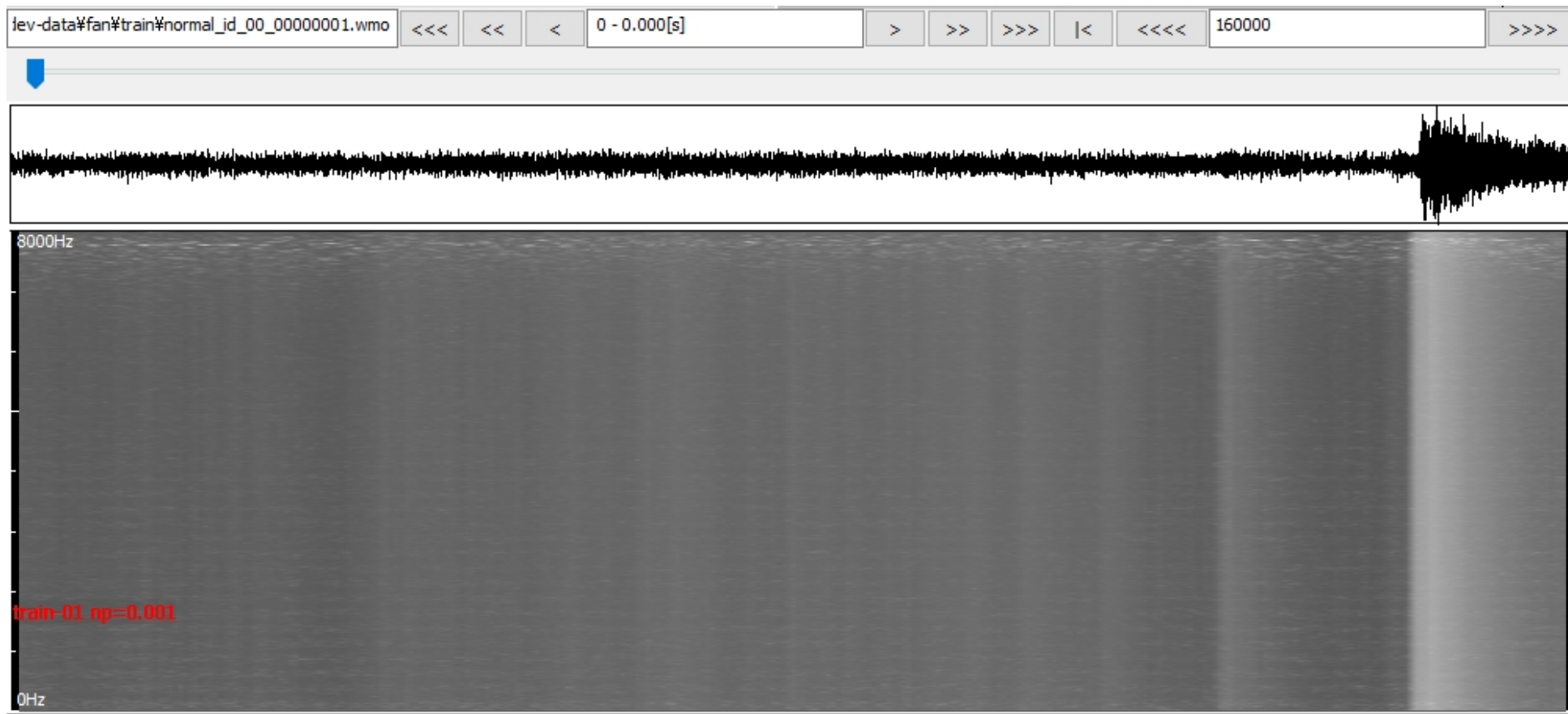


図 1-2-1-rst. normal_id_00_00000001.wav に対する
「局所的位相シフト検出法」によるスペクトログラム (0Hz-8000Hz)

図 1-1-x-rst.と図 1-2-x-rst.を目視で見ると、異なる試行であっても「局所的位相シフト検出法」によるスペクトログラムは安定している。共通して目立つ特徴としては、1000Hz の間におよそ 4 本の割合で見える白く横向きに走るバンドが挙げられる。またバンド上の白は一様に白くはなく、縦方向に雑音を周波数表示したような特徴を有している。

また図示はしないが、DCASE2020 の fan に用意されてる「id_00」「id_02」「id_04」「id_06」のそれぞれに対して先頭の 5 個のデータを同様に確認したところ（合計で 20 個=4×5）、やはり「局所的位相シフト検出法」によるスペクトログラムは安定していて白く横向きに走るバンドが見える。

ただし、「0.はじめに」で示した

URL: <https://sigprocrandwalk.hatenadiary.org/entry/20170722/1500751760>

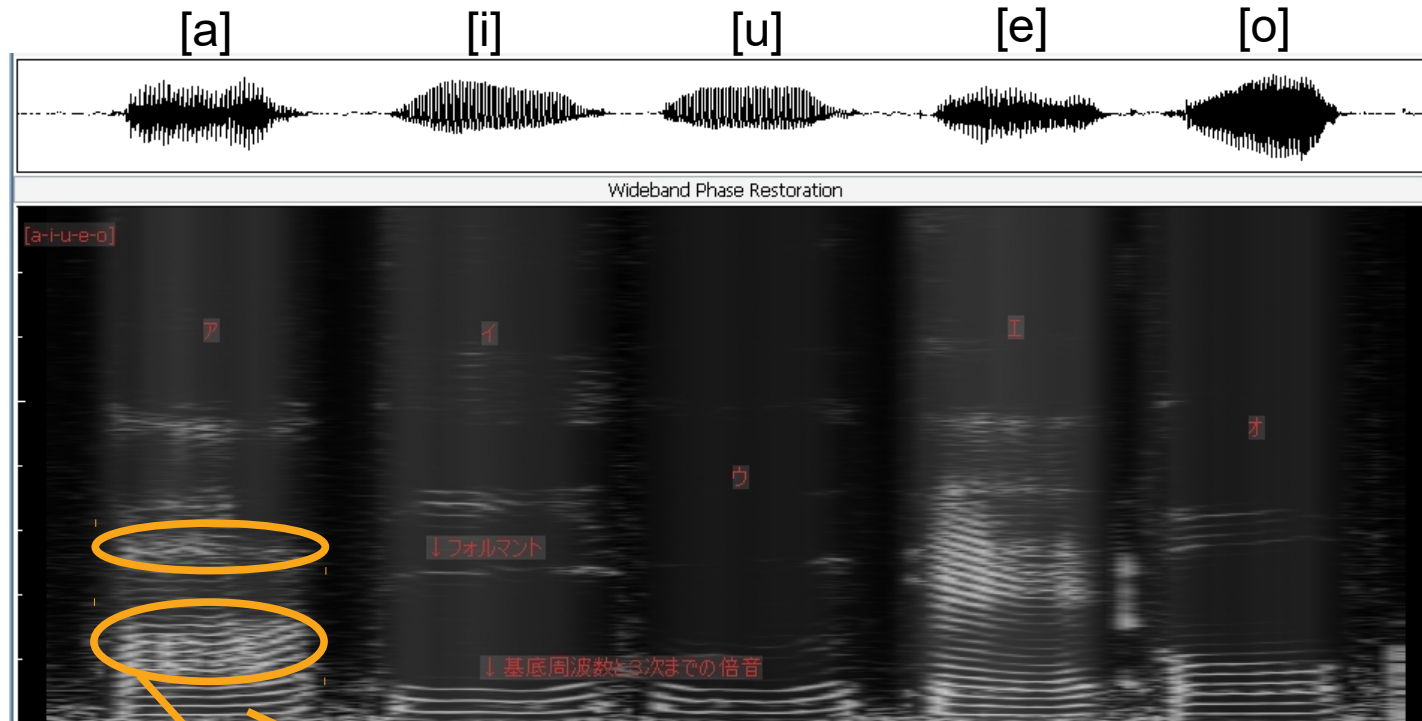
URL: <https://sigprocrandwalk.hatenadiary.org/entry/2020/03/22/193447>

では、「局所的位相シフト検出法」によるスペクトログラムは正しいことを示せたが、DCASE2020 のデータについては実測データであることから正解が判らず、正しいかどうかは簡単には言えない。

そこで、次の章では音声解析で用いる「フォルマント」と似た考え方を用いて、図 1-1-x-rst.と図 1-2-x-rst.で得られている結果について説明を試みる。

2. 「DCASE2020」の「task2」の「fan」に対して「局所的位相シフト検出法」で得られた結果の考察

2-1. 音声解析で用いられる「フォルマント」という考え方



音源である声帯の基底周波数（+その高調波で細かい縞に見える）

フォルマント：共鳴器（声道+口腔）の形状で決まる共鳴周波数による強調

図 2-1. 音声解析におけるフォルマントの一例

フォルマントとは、音源の周波数特性が共鳴器の共鳴周波数（固有振動数）により部分的に強められたり弱められたりした結果、スペクトログラム上に表れる模様のことである。音声解析においては、有声音（主に母音）の場合、音源は「声帯」で共鳴器は「声道＋口腔」と考えられる。「声道＋口腔」の共鳴周波数はその形状により決まるが、例えば母音である[a]、[i]、[u]、[e]、[o]を発する場合、それぞれに応じた口腔の開け方があり、フォルマントが生じる周波数を調べることで、母音の[a]、[i]、[u]、[e]、[o]を判定できる。

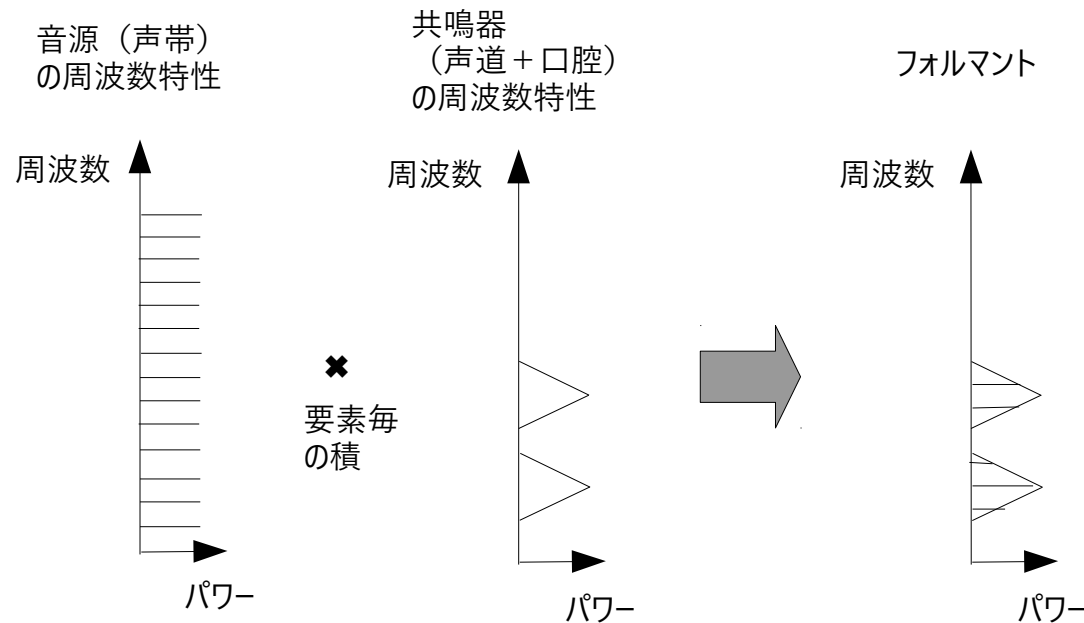


図 2-2. フォルマントができる様子

また、実際に検討されている例を見たことはないが理屈の上では、無声音（無声子音）にもフォルマントは存在する。例えば[s]の場合、音源は「歯間を流れる空気の風切り音（雑音）」で、共鳴器は「声道＋口腔」と考えることができる。ただし[s]を発する際の「上下の歯を合わせて空気を歯間から押し出す」という動作の関係上、口腔の形状に自由度は無くフォルマントはほぼ固定されることになる。

2-2. 「DCASE2020」の「fan」に対するフォルマントという考え方の導入

「DCASE2020」の「fan」にフォルマントという考え方を導入する場合、

音源： fan の風切り音（雑音）

共鳴器： 筐体を含む fan 全体（筐体を含む fan 全体の形状で固有振動数が決まる）

とするのが適切と考えられる。

以上の話の妥当性を検証するには、まず

- 1) DCASE2020 で用いられた fan の共鳴周波数（固有振動数）を実測する
（物体の共鳴周波数を実測する方法は工業的に確立されている）
- 2) fan の形状（詳細な形状と実寸）と材質を基に共鳴周波数を計算する

などの方法で共鳴周波数を得る。

更に音源と共鳴器の情報を基にフォルマントを計算するが、図 1-1-x-rst.と図 1-2-x-rst.で得られている結果と合致すれば、間接的に「局所的位相シフト検出法」は正しいと推測できたことになる。

以上の流れについて、今後は可能なら 1) もしくは 2) に取り組み、DCASE2020 の fan に対して「局所的位相シフト検出法」で正解が得られてるかを明らかにしたい。

最後に、DCASE2020 はより良い機械学習の方法を競うためのチャレンジであるが、機械学習や統計的分類手法では、「座標変換」や「写像」が「『観測信号』から『原信号』を回復するためのもの」である必要は無く、「分類対象（DCASE2020 の fan の場合は「normal」と「anomaly」）を分けるのに役立てば良い」という考え方がある。次の章ではこの章の話題（DCASE2020 の fan に対して「局所的位相シフト検出法」で正しい結果が得られてるか）とは別に、「局所的位相シフト検出法」を用いることで分類が容易になるかどうかについて述べたい。

3. 「局所的位相シフト検出法」が機械学習及び統計的分類手法に役立つ可能性について

1. では分類対象の片方である「normal」に関するスペクトログラムを示したが、この章ではもう片方である「anomaly」を同様に示し、両者を比較できるようにする。

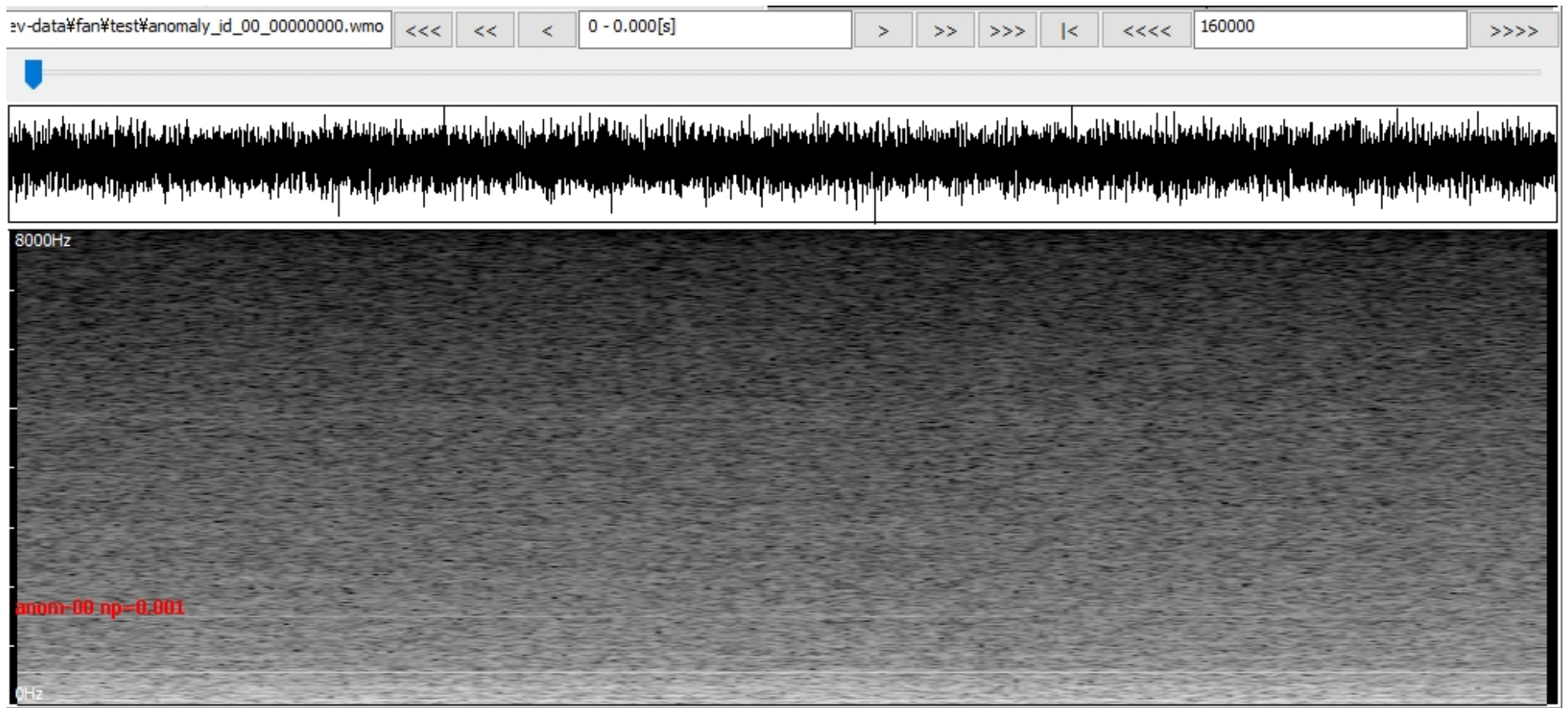


図 3-1-1. anomaly_id_00_00000000.wav に対する
「従来方法」によるスペクトログラム (0Hz-8000Hz)

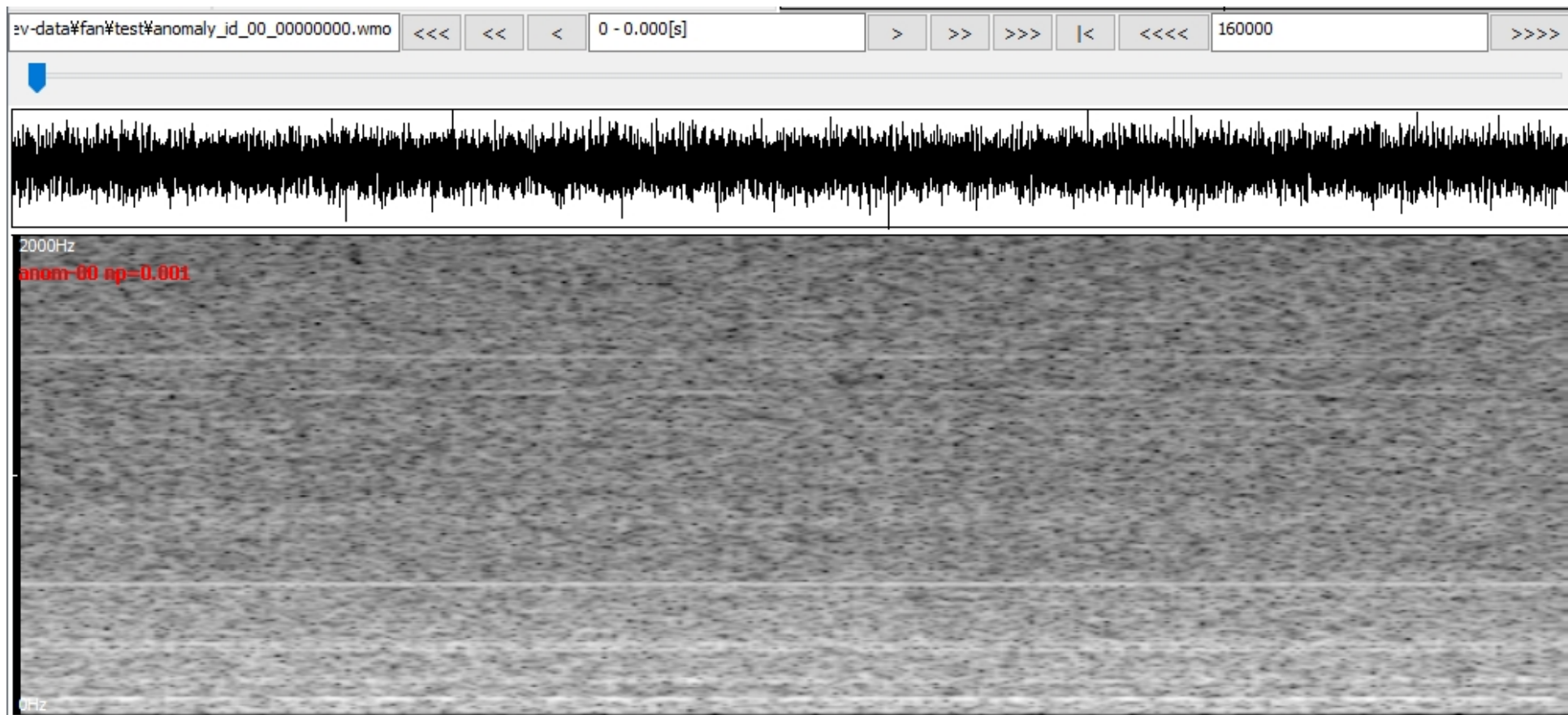


図 3-1-2. anomaly_id_00_00000000.wav に対する
「従来方法」によるスペクトログラム (0Hz-2000Hz)

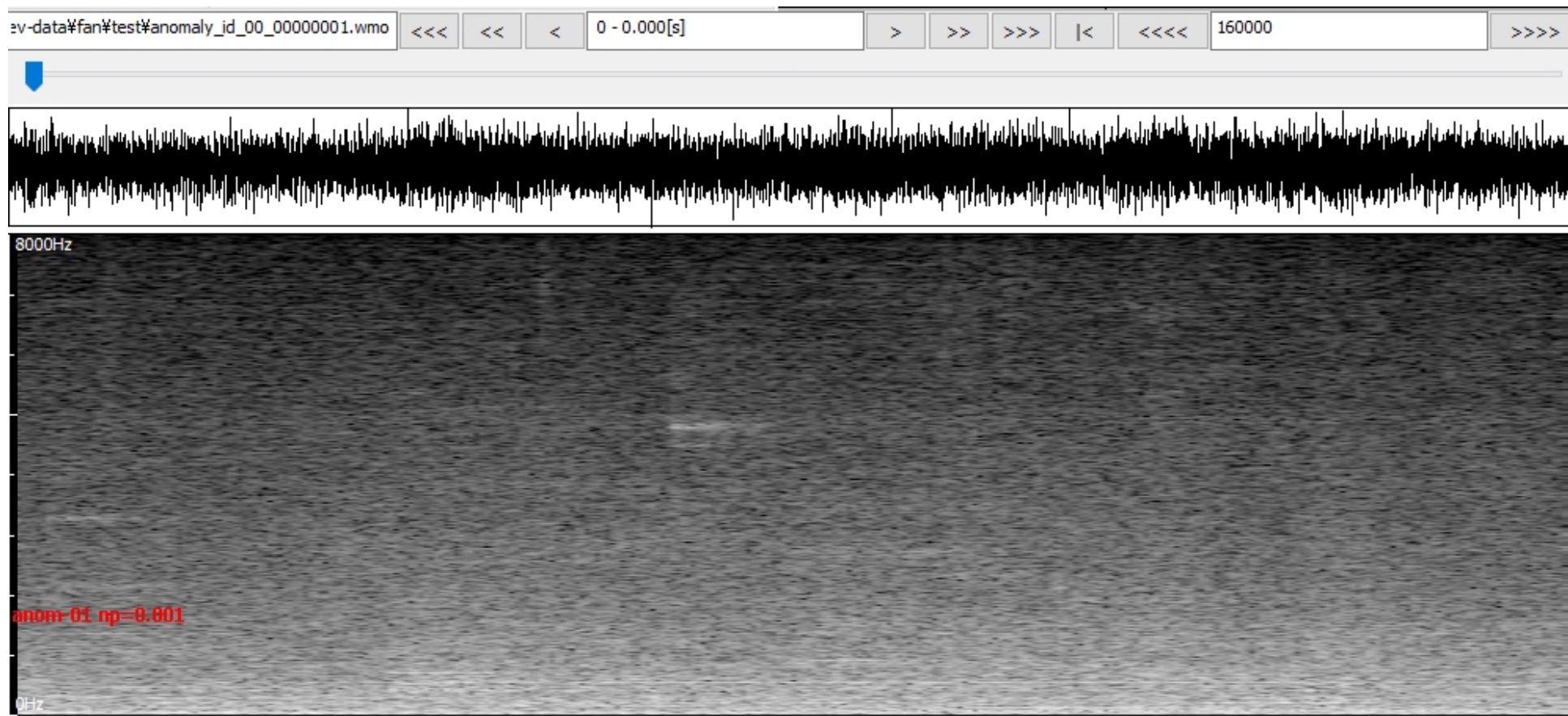


図 3-2-1. anomaly_id_00_00000001.wav に対する
「従来方法」によるスペクトログラム (0Hz-8000Hz)

データに付けられたファイル名（anomaly_id_00_00000000.wav など）のサフィックスは次の意味を表している。

左：anomaly（normalかanomalyのanomaly）

中：id_00（複数あるデバイス（fan）のデバイス番号）

右：00000000（同一デバイスに対して複数試行を行った際の番号）

以上のサフィックスより、図 3-1-x.と図 3-2-x.は同一デバイスにおける異なる試行番号の結果を表している。図 3-1-x.と図 3-2-x.を目視で見ると、「従来方法」によるスペクトログラムでは、同一デバイスにも拘わらず結果が安定して見えない。ここが安定していない場合、続けて処理（統計的分類や機械学習など）を行う場合に不利になる可能性がある。

以下では、図 3-1-x.と図 3-2-x.と同一のデータに対して、「局所的位相シフト検出法」を用いた場合のスペクトログラムを示す。

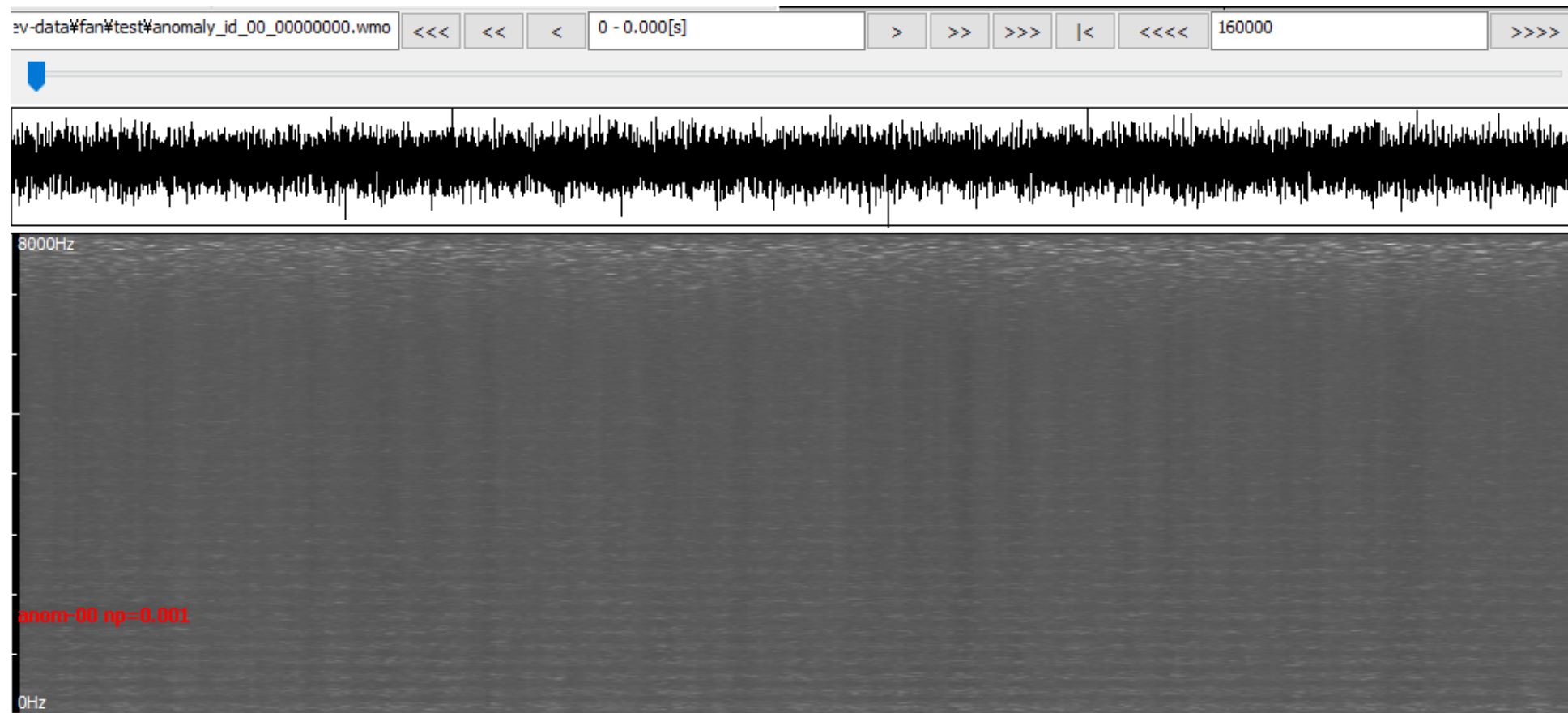


図 3-1-1-rst. anomaly_id_00_00000000.wav に対する
「局所的位相シフト検出法」によるスペクトログラム (0Hz-8000Hz)

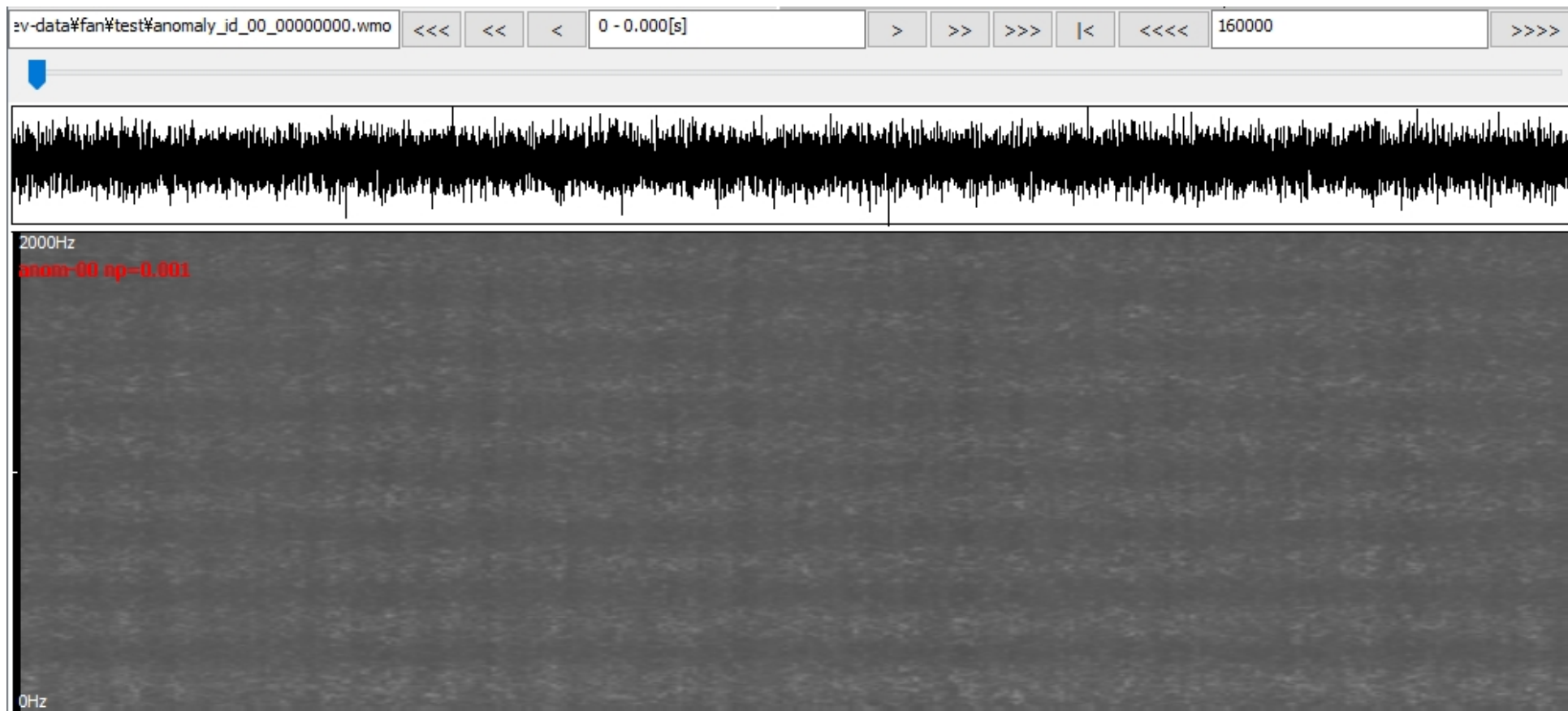


図 3-1-2-rst. anomaly_id_00_00000000.wav に対する
「局所的位相シフト検出法」によるスペクトログラム (0Hz-2000Hz)

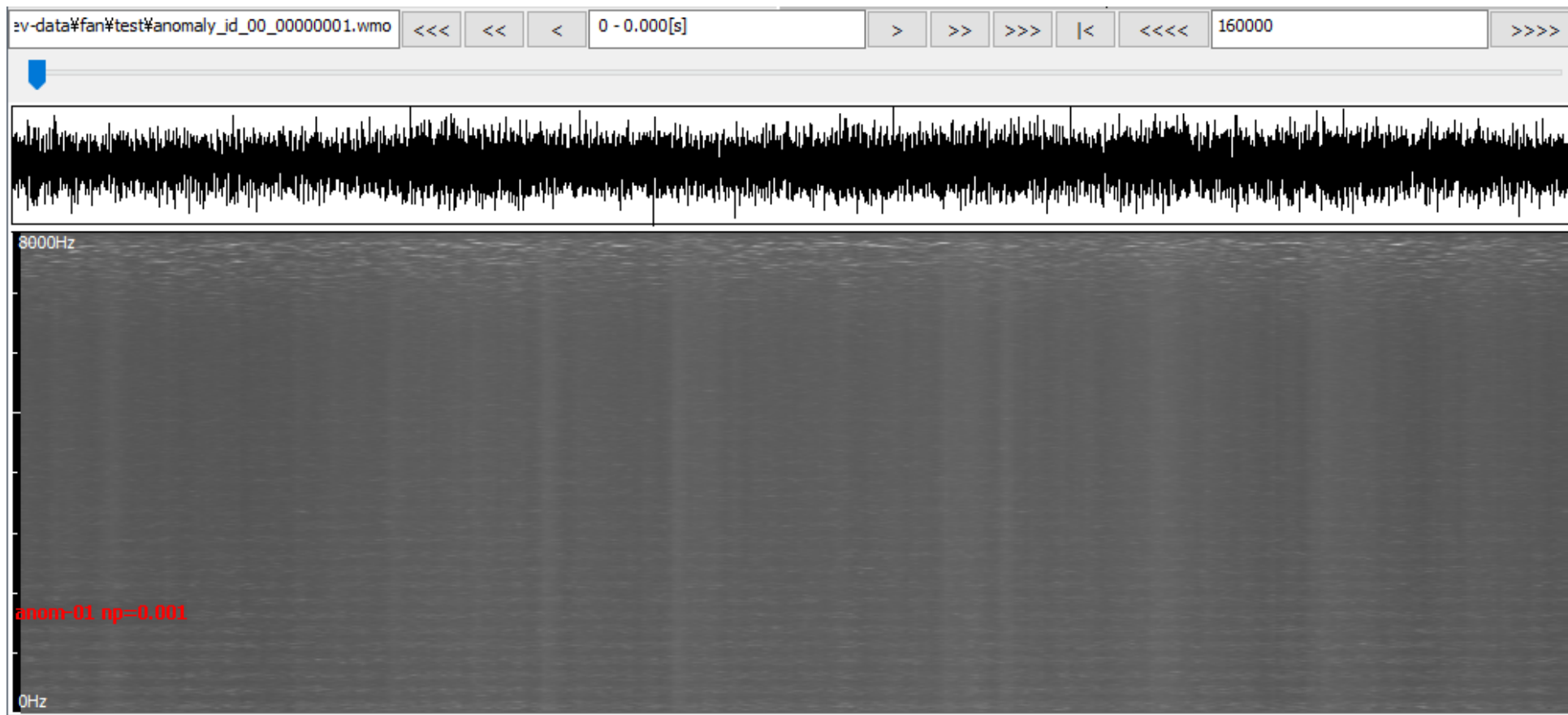


図 3-2-1-rst. anomaly_id_00_00000001.wav に対する
「局所的位相シフト検出法」によるスペクトログラム (0Hz-8000Hz)

図 3-1-x-rst.と図 3-2-x-rst.を目視で見ると、異なる試行であっても「局所的位相シフト検出法」によるスペクトログラムは安定している。共通して目立つ特徴としては、1000Hz の間におよそ 4 本の割合で見える白く横向きに走るバンドが挙げられる。またバンド上の白は一様に白くはなく、縦方向に雑音を周波数表示したような特徴を有している。

また図示はしないが、DCASE2020 の fan に用意されてる「id_00」「id_02」「id_04」「id_06」のそれぞれに対して先頭の 5 個のデータを同様に確認したところ（合計で 20 個=4×5）、やはり「局所的位相シフト検出法」によるスペクトログラムは安定していて白く横向きに走るバンドが見えることが判った。

normal を示してる「図 1-1-x-rst.と図 1-2-x-rst.」と anomaly を示してる「図 3-1-x-rst.と図 3-2-x-rst.」を目視で比較すると、

- 1) 「1000Hz の間におよそ 4 本の割合で見える白く横向きに走るバンド」が存在するという点については normal も anomaly も同じである
- 2) 1) のバンドについて、normal より anomaly の方がコントラストが高く、また高周波領域まで特徴が伸びている

また図示はしないが、DCASE2020 の fan に用意されてる「id_00」「id_02」「id_04」「id_06」のそれぞれに対して先頭の 5 個のデータを同様に確認したところ（合計で 40 個、normal 分 20 個と anomaly 分 20 個）、上記の 1) と 2) について特徴として同様であった。

「fan」は工業製品であることから、例えば normal と anomaly の違いがあったとしても全体として似た状態になるのは考えられることであり、1) はおかしいこととは言えない。そして normal と anomaly の違いは 2) に表れてると言える。

今のところはいくまでも目視での確認であるが、今後は 1) と 2) の特徴を評価関数化し、機械学習もしくは統計的分類手法に組み込んで性能を評価したい。

4. おわりに

もし「他のデータについて試したい」というご希望があればメールにてお知らせください。

園部和夫 [aka SigProcRandWalk | SPRW , mailto : fuja2_at_bj8.so-net.ne.jp]